THE INTRADAY PATTERN OF INFORMATION ASYMMETRY: EVIDENCE FROM THE NYSE

A Thesis Submitted to
The College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Finance
in the Edwards School of Business
University of Saskatchewan
Saskatoon, Saskatchewan

By
Juan Wang

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Degree of Master of Science in Finance from the University of Saskatchewan, I agree that the Libraries of the University of Saskatchewan may make it freely available for inspection. I further agree that permission for copying this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or make other use of materials in this thesis in whole or in part should be addressed to:

Head of the Department of Finance and Management Science

Edwards School of Business

University of Saskatchewan

Saskatoon, Saskatchewan

S7N 5A7

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my co-supervisors, George Tannous and Craig Wilson, for giving me an opportunity to write my thesis under their direction. Their guidance, assistance, support and encouragement were imperative for the creation of the thesis. I also would like to thank my internal examiner, Dr. Marie Racine, and my external examiner, Dr. Gordon Sick, for their invaluable suggestions and comments.

I would like to give my appreciation to the Edwards School of Business Technology Support Center for their assistance. My thanks also go to all the other faculty and staff of Master of Science in Finance Program for their support during my studies in this program. I also would like to thank the College of Graduate Studies and Research for awarding me the university graduate scholarship.

Lastly, I would like to thank my parents and friends for their love and help in my life.

ABSTRACT

Previous studies (e.g. Benston and Hagerman, 1974, Bagehot, 1971 and Stoll, 1978) suggest that the bid-ask spread consists of three components: asymmetric information cost, inventory holding cost, and order processing cost. Other literature (e.g. Brock and Kleidon, 1992, Heflin et al, 2007, and McInish and Van Ness, 2002) reports that the bid-ask spread varies during a trading day following a U-shaped pattern. One explanation for this observation is that it is the result of changes in information asymmetry costs over the trading hours, assuming the other costs are fixed. However, no empirical study directly measures how information asymmetry changes over the trading day. We explore how this measure relates to the spread as well as the quote depth.

Our research divides a trading day into 13 half-hour trading intervals and measures information asymmetry during each interval following the model developed by Madhavan and Smidt (1991) and Noronha et al (1996). Their model can directly estimate the level of information asymmetry in each interval. This enables us to observe the intraday pattern of information asymmetry directly and compare it to the patterns of the spread and the quote depth. Furthermore, we test the relationship between the spread and the information asymmetry and the relationship between the depth and the information asymmetry in a dynamic context to see how market makers manage information risk over trading hours.

We find that the risk of information asymmetry varies significantly during the trading day. There is a large drop over the first interval, and another large drop over the last interval, with smaller fluctuations over the remaining intervals. Moreover, we show that the spread is consistent with an L-shaped pattern as opposed to the U-shaped pattern proposed by previous studies while the depth is increasing throughout the 13 trading intervals. Furthermore, we observe that the variations of the spread and the depth are respectively positively and negatively related to the intraday variations in the degree of information asymmetry across the trading intervals. In particular, a large decline in information asymmetry at the beginning of the day is associated with a large reduction in the spread, whereas a large decline in information asymmetry at the end of the day is associated with a large increase in the quote depth. This emphasises the importance of studying both measures of liquidity simultaneously.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Market makers are intermediaries in the security market, who provide liquidity by using their own account to take the role of counterparty for submitted market orders. They post the highest price at which they would like to buy the security and the lowest price at which they would like to sell the same security. These quoted prices are respectively called the bid and the ask quotes and the difference between the bid and ask quotes is called the bid-ask spread (hereafter denoted simply as spread). Market makers also disclose the number of shares they are willing to buy and sell at their quoted prices. The number of shares available for trading provided by market makers at the bid price and ask price are respectively called bid depth and ask depth (summarised as quote depth or simply depth). The ask price is generally strictly higher than the bid price. Thus, the spread is an important component of the transaction costs faced by investors. On the other hand, the spread represents an important component of a market maker's revenue.

The spread consists of three components: order processing cost (OP), inventory holding cost (IH), and asymmetric information cost (AI). Benston and Hagerman (1974) and Stoll (1978) indicated that the OP cost involves the direct costs of arranging trades, matching orders, and recording transactions. Stoll (1978) and Tinic (1972) described inventory holding costs as those associated with holding diversifiable risk. Bagehot (1971) indicated that the asymmetric information cost arises through an adverse selection problem.

Previous studies of the NYSE showed that the spread quoted by market makers exhibits a U-shaped pattern within a trading day. The spread is at the highest level at both the opening and closing of a trading day (Brock and Kleidon, 1992, Foster and Viswanathan, 1993, Heflin et al, 2007, Lee et al, 1993, Madhavan et al, 1997, McInish and Wood, 1992). This intraday pattern of spread implies that investors are facing higher trading cost at the open and close of a trading day. There are three explanations for the intraday pattern of spread: information asymmetry risk effect (Foster and Viswanathan, 1994 and Madhavan, 1992), inventory control effect (e.g. Amihud and Mendelson, 1982), and market maker power effect (Block and Kleidon, 1992).

Compared to the spread, fewer studies focus on the depth quoted by market makers.

Dupont (2000) and Lee et al (1993) pointed out that the spread and the depth are two main indicators of the market liquidity and market makers simultaneously increase the spread and lower the depth to manage the information risk. However, Lee et al (1993) link the changes of the spread and the depth to the change in degree of information asymmetry by testing the changes of the spread and the depth around earnings announcements. Few prior studies tried to integrate the intraday variation of the spread and the depth and the intraday variation of the information asymmetry. Since the spread and the depth quoted by market makers represent the level of liquidity in the market, and liquidity has implications for market efficiency, it is important to understand how the spread and the depth change and how the information asymmetry affects the quoting behaviour of market makers over different trading hours.

In this study, we use transactional data from the NYSE to examine the intraday pattern of the information asymmetry based on the direct estimation of information asymmetry using the method developed by Madhavan and Smidt (1991). We also examine the intraday patterns of the spread and the quote depth. We find that information asymmetry is at the highest level at the beginning of a trading day, then declines until the noon period and it increases slightly after noon and then it drops again toward the end of a trading day. In particular, it drops substantially in the opening and closing intervals, with smaller changes occurring over the remaining intervals.

We find that the intraday pattern of the spread is considerably different from the information asymmetry pattern, particularly at the closing time. The spread contracts substantially after the opening interval; however, it does not change obviously at the end of the trading day. We suspect this behaviour is due to the monopoly effect. Moreover, the depth is found to be lowest at the open and it increases consistently toward the close, although the changes are relatively small until the last interval, which involves a substantial increase. Furthermore, we empirically examine the relationship between information asymmetry and the spread as well as the depth in a dynamic context. We confirm that market makers quote a higher spread and lower number of shares for trading while they detect higher informed trading in the market.

Collectively, this study reveals the intraday patterns of information asymmetry, the spread and the depth. We also find that the degree of information asymmetry is a key factor affecting the market makers' quoting behaviour. Firstly, since the spread is part of the trading costs faced by investors and the spread combined with the depth are factors relating to the liquidity in the market, our study offers useful information to investors and financial professionals. For investors and professionals, they need to know the information of trading costs while they are evaluating a proposed trading strategy. Particularly, understanding how the overall liquidity is changing over the trading hours will be useful to traders who are attempting to liquidate significant positions. Secondly, in order to maintain a fair and efficient stock market, a fundamental factor of the market considered by policy makers is the provision of liquidity. A highly liquid market will encourage more investors to participate in trading, (or equivalently, too wide spread and too low depth might drive investors away) which can lower the cost of raising capital and encourage economic growth. Therefore, policy makers should better understand market makers' quoting behaviour for the purpose of building a fairly liquid stock market. Thirdly, our study helps academics to understand better the information flow of stock markets, and market makers' quoting behaviour.

Chapter 2: Literature Review

2.1 Components of the Spread

According to previous research, a market maker faces three main costs of doing business. (1) The direct costs are called order processing costs (OP). These are primarily fixed costs that are allocated to each transaction. Benston and Hagerman (1974) and Stoll (1978) indicated that the OP cost involves the costs of arranging trades, matching orders, and recording transactions, etc. (2) Stoll (1978) and Tinic (1972) described inventory holding costs (IH) as those associated with holding diversifiable risk. A market maker must maintain an inventory that is sufficient to provide liquidity to the market even if there is significant buying pressure. This naturally forces the market maker to hold an undiversified portfolio. The cost can be determined by the expected difference in revenue from holding a well-diversified portfolio of equivalent risk, allocated to each transaction. (3) The third component, called asymmetric information cost (AI), was first indicated by Bagehot (1971). This cost arises through an adverse selection problem. Since market makers generally have less information about the true value of a security than do company insiders or other informed traders, the market maker would be expected to lose money by trading with such informed traders. As the trading process is generally anonymous (or can be handled through a third party), by quoting bid and ask prices, the market maker inadvertently attracts such informed traders whenever her quoted prices diverge from the informed traders' perception of the true value, creating a moral hazard problem. The asymmetric information cost, therefore, represents the expected loss to informed traders per transaction. Glosten and Milgrom (1985) prove that a bid-ask spread can arise in an asymmetric information setting even when other transaction costs are zero and market makers are risk neutral and perfectly competitive.

In order to remain solvent, the market maker's revenue must cover these costs. This gives a natural decomposition of the spread into the three cost components (as measured on a per transaction basis). Of course, this decomposition ignores the market maker's potential profit, which is an important consideration for investors trading in a market with limited market maker competition such as the NYSE.

2.2 Empirical Tests for the Components of Spread

Following the theoretical analysis of the three determinants of the spread, many researchers estimated the magnitude of the three components using various methods. Roll (1984) modeled the serial covariance of the transaction price to measure the effective spread. He argued that the observed negative serial dependence of the transaction price can be attributed to the existence of the market maker's spread. However, his model was designed to analyze the size of the OP costs. Following Roll (1984), Choi et al (1988) extended the model to allow the dependence of the transaction type which was assumed by Roll (1984) to have equal probability to be a buy or sell order at every trading time. However they did not attribute the serial dependence of trading type to the effect of inventory cost, but rather argued that it is caused by broken orders and limit orders. Stoll (1989) argued that under the inventory control model, the probability of trade reversal (a buy order following a sell order or vice versa) exceeds 0.5. In addition, he took the order processing cost and asymmetric information cost into consideration, and decomposed the three components based on the properties of two time series: the serial covariance of transaction returns and the serial covariance of quoted returns. His empirical results, based on data from the NASDAQ traded securities, showed that the OP, IH and AI costs are respectively 47%, 10% and 43% of the total costs. George et al (1991) indicated that previous estimation of the components based on the covariance of transaction price or return assumed that the expected return stays constant over the modeling period, which would not be true in such a setting. Therefore, the previous estimated results could be biased. They solved this problem by using the covariance of the difference between the transaction price based return and quote price based return. Their results show that the OP cost is the most significant component of the spread. The adverse selection cost turned out to comprise 8% and 13% respectively based on daily and weekly data. Although significant, they pointed out that the proportion they derived for AI cost will only be appropriate under the assumption that the spread is independent of trade size. However, according to Easley and O'Hara (1987), the AI component should increase with trade size. In this case, George et al (1991) argued that their estimate at least provided the lower bound of the AI component for large order size transactions.

In the studies indicated earlier, empirical measurement of the components is based on the covariance properties of transaction prices. In another strand of research, the spread decomposition is based on the trade indicator regression model. Glosten and Harris (1988) and Hasbrouck (1988) argued that the properties of order flow can uncover information related to the quote revision, which could show distinctive features of the IH and AI components. Therefore, they included the trade indicator into the regression model to measure the components of the spread. Glosten and Harris's (1988) empirical result showed that the asymmetric information cost explains approximately 1/3 of the spread. Hasbrouk (1988) also found that the AI component to be substantial. Madhavan et al (1997) took the innovation or unexpected order flow into account and constructed a trade indicator model to measure the adverse-selection component. Moreover their test was based on intraday transaction data. However, all of these methods failed to distinguish between the other two components, OP cost and IH cost, which are estimated together as a transitory component. Huang and Stoll (1997) extended the trade indicator model and successfully separated the IH and OP components.

From previous theoretical and empirical studies, it is generally accepted that the AI cost constitutes a significant component of the spread. This means that when a high degree of information asymmetry is detected, market makers tend to increase the AI component of the spread as a reaction to increased risk due to information arrival. McInish and Wood (1992) examined the relationship between the spread and the amount of information coming to the market and found them to be positively related. The variable they used in their model to measure the information generated by the market is the normalized trading size that will "capture the effect of unusually large or small trades relative to the average size of trades" (p758). Gregoriou et al (2005) also empirically demonstrate that market makers widen the spread when they are at an informational disadvantage to informed traders. They measure the degree of information asymmetry by the level of disagreement in analysts' earnings forecasts. Moreover, Gong (2007) also found the spread to be positively related to the degree of information asymmetry.

2.3 Intraday Pattern of Spread and its Components

Since the early 1990's, the availability of transaction-by-transaction data has allowed a growing body of researchers to explore the intraday pattern of the spread and its components.

2.3.1 Evidence from NYSE:

On the NYSE, specialists, viewed as market makers, are dealers representing a NYSE specialist firm. Specialists provide liquidity when there's a demand-supply imbalance by purchasing to or selling out of their own inventory to equalize and stabilize the market. The bid-ask spread is the price they charge to offer this liquidity to investors.

Brock and Kleidon (1992) observed a wide spread at both the open and close of trading on the NYSE. Based on the quotation records from specialists in 1989, McInish and Wood (1992) found that the spread is highest at the third minute of the opening hour, then declines over the trading hours, and increases slightly toward the close of trading. This reverse J-shaped pattern was illustrated by plotting the minute-by-minute average spread across stocks, which is more consistent with the L-shaped pattern we document. Furthermore, their estimated parameters of dummy variables for different trading intervals during the day also showed a consistent reverse J-shaped pattern. Foster and Viswanathan (1993) inferred the intraday pattern of trading cost by separating the intraday pattern of AI cost from the remainder of the spread. Their empirical result showed that the AI cost is high at the first half hour of a trading day, declines during the midday, then rises toward the close. In contrast, they find little intraday variation in the remaining part of the spread, referred to by the authors as fixed cost. Evidence from Lee et al (1993) also showed a U-shaped pattern of the intraday spread. Madhavan et al (1997) derived a model to decompose the spread, and then demonstrated the U-shaped intraday pattern of spread and found that the AI component decreases over the day and levels off near the close.

The NYSE is an example of a hybrid quoting system, in which both specialists and limit order traders provide liquidity to investors. Under this system, limit orders are expected to play an important role in establishing security prices and affecting bid-ask spreads. All of the previously mentioned studies included all bid and ask quotes for NYSE stocks, reflecting both limit-order interest and specialist interest. Chung et al (1999) pointed out that on the NYSE, limit order traders, unlike specialists, can post either the bid or ask price and trade only on one side. When the limit order price is better, specialists must reflect the highest bid price and the lowest ask price from those limit orders. They reported that around 74.9% of quotes originate from limit orders on at least one side, and they conclude that the U-shaped pattern derived from the whole sample is mainly driven by the limit orders. The specialist component of the spread is highest at the opening hours, then narrows and doesn't show a rise toward the end. Heflin et al (2007) considered firms with high and low disclosure quality and found that the intraday spread exhibits a U-shaped pattern for the two groups.

McInish and Van Ness (2002) obtained intraday trades and quotes data from the TAQ database from the NYSE, which contains quotes from both limit-order traders and specialists, to examine the intraday patterns for the spread and the components. Using the decomposition model of George et al (1991) and Madhaven et al (1997), they show that the spread follows crude U-shaped pattern and the AI component follows the reverse J shape.

2.3.2 Evidence from NASDAQ and Comparison between NYSE and NASDAQ

Besides the intraday studies based on NYSE data, there is another strand of related research examining the intraday spread of stocks traded on NASDAQ and comparing the results from the two stock markets. According to Van Ness et al (2002), the main difference between the NYSE and the NASDAQ in the market making process is the single specialist system on the NYSE versus multiple dealers on the NASDAQ. On the NYSE, each stock is assigned to one specialist while on the NASDAQ each stock is usually actively traded by several dealers.

Chan et al (1995) showed that the spread for NASDAQ stocks is highest during the opening half hour, then declines slowly and continues to stay stable until the close when a sharp decline occurs. This pattern is different from the U-shaped pattern reported for the NYSE. Their empirical tests also indicated that the different pattern of spread between the two stock markets is not explained by the trading volume and trading volatility which are comparable between NYSE and NASDAQ. Therefore, they attribute the different patterns of spread to the distinct institutional characteristics under the assumption that these two markets have similar information flow. However, they didn't compare the pattern of information flow among these two markets. Chung and Van Ness (2001) and Chung and Zhao (2003) argued that if the different order handling rules (OHR) brought about the different spread pattern between the NYSE and the NASDAQ, the two markets should exhibit similar spread pattern after the implementation of the new OHR on the NASDAQ in 1997. This change allowed limit orders to compete with dealer quotes as on the NYSE. Undertaking empirical tests with a sample period just following the new OHR implementation, Chung and Van Ness (2001) failed to see such convergence. Chung and Zhao (2003) tested the same hypothesis again using data two years after the new OHR implementation, and found that the intraday pattern of NASDAQ spread does exhibit a U-shaped pattern similar to that shown for the NYSE spread. They explained that the different results may be due to the market assimilation of the new OHR. However, Ascioglu et al (2006), using a data sample during 2005, showed that the spread of stocks traded both on NYSE and NASDAQ "fall in a marked way after the first half hour of trading, and remain fairly flat throughout the rest of the day suggesting that the dramatic changes in the marketplace during the last decade have altered the intraday pattern of spreads for NYSE and NASDAQ stocks" (p2).[1]

Regarding the relative magnitude of the AI component between the NYSE and the NASDAQ, Affleck-Graves et al (1994), using a matched sample of NYSE and NASDAQ stocks, empirically showed that the AI component of the NYSE spread is significantly higher than that observed on the NASDAQ. However, Van Ness et al (2002) showed contrary results with NASDAQ having the higher AI component. They ascribe these contrary results to the different spread decomposition models used by the different authors. Ascioglu et al (2006) showed that for both markets, the AI component is at its widest level during the opening hour. However, they do not estimate the information flow in the market directly. Instead, they analyse the AI component derived from the spread, employing the decomposition model of Madhavan et al (1997). This AI component can be viewed as an indirect measurement of the information asymmetry since as we have mentioned above, market makers tend to increase the AI cost of the spread when they detect higher information asymmetry. However, the spread is not the only way for market makers to manage information risk. Instead, market makers can also choose to reduce the quote depth when they detect a rise in information asymmetry. Therefore, this estimate may not accurately measure information asymmetry.

[1] Most previous intraday research on spreads in both NYSE and NASDAQ were based on data samples obtained from periods before 2000.

2.4 Explanations of the Intraday Pattern of the Spread

Historically, there are three arguments proposed to explain the intraday spread pattern: market maker power, inventory control, and adverse selection. Brock and Kleidon (1992) argued that the wide spread during the opening and closing hours is due to market makers' privileged knowledge of order imbalance and their monopolist power over investors' inelastic demand. Amihud and Mendelson (1982) suggest that market makers widen the spread as closing approaches to control their inventory before the overnight non-trading interval. They argue that during closed hours there is limited opportunity for market makers to adjust their inventory imbalances. In contrast, the proponents of the information asymmetry theory, for example Madhavan (1992), argued that trades accumulate over the trading hours, private information is gradually impounded into the trading price, and a higher level of private information is expected during the opening hour. Kyle (1985, p1316) also indicated that "The informed trader trades in such a way that his private information is incorporated into prices gradually." This environment would lead market makers to increase their spread in order to mitigate such risk. Foster and Viswanathan (1994) also argued that the wider spread can be caused by the competition between two groups of differently informed traders.

2.5 Relationship between the Depth and the Information Asymmetry, and Intraday Pattern of the Depth

Quote depth refers to the number of shares available for trading at the bid and ask prices, reflected in each quotation. Quotes from market makers contain the bid and ask prices combined with the corresponding depths. Only a few studies are dedicated to the examination of depth, especially with respect to its intraday pattern. Lee et al (1993) argued that both the spread and the depth are important dimensions of liquidity. The spread reflects the price dimension while the depth quoted by market makers reflects the quantity dimension. When more informed traders place orders in the market, market makers can increase the spread or reduce the quantity of shares available for trading to guard against information risk. Either activity has the effect of increasing the price impact of trades. Their empirical results show a negative relationship between the spread and depth around earnings disclosure events and also on an intraday basis. This finding indicates that market makers use both the spread and the depth to manage information risk. Dupont (2000) supports this conclusion by arguing that the quote depth can be used by market makers to fend off the information risk that arises from increased informed trading. He showed that the depth quoted by market makers is more sensitive to changes in the degree of information asymmetry than the spread. Concerning the intraday pattern of quote depths, Li et al (2005) separately tested the patterns of depth quoted exclusively from specialists and limit orders from the NYSE and found an inverted U-shaped pattern for both.

2.6 Summary of the Literature

Table 1 summarizes the literature mentioned above.

======Insert Table 1 Here======

Previous studies evidenced two different intraday patterns of the spread. Most studies agreed that for the NYSE the intraday pattern of the spread is U shaped whereas others (Ascioglu et al, 2006 and McInish and Wood, 1992) document a reverse J-shaped intraday pattern more consistent with our finding. Researchers tried to explain the intraday pattern of the spread in terms of information asymmetry (Foster and Viswanathan, 1994 and Madhavan, 1992). They report that the AI component of the spread is roughly decreasing over the trading hours (Foster and Viswanathan, 1993, Madhavan et al, 1997 and McInish and Van Ness, 2002). They focused on the AI cost decomposed from the spread. However, the AI component derived from the spread might not be a good proxy for the level of information asymmetry, because market makers manage information risk both by changing the quoted spread and by changing the quoted depth. Since the AI component of the spread might not be an exact measurement of the level of information asymmetry, it is more appropriate to see how the spread (containing the AI component) responds to changes in information asymmetry.

Therefore, in our study, we estimate the degree of information asymmetry over trading hours based on a model adapted from Madhavan and Smidt (1991) and Noronha et al (1996). This model provides a way to directly estimate the level of information asymmetry across the different time intervals within a trading day. The first objective of our study is to show how information asymmetry changes within a trading day. In addition, we show the intraday pattern of the spread and the depth and test whether the information asymmetry factor can explain the intraday behaviour of these two.

Secondly, previous studies, for example Gong (2007) and Gregoriou et al (2005) have argued that market makers are likely to widen the spread while they are at an information disadvantage. However their conclusions are based on cross sectional tests which show that stocks with a higher degree of information asymmetry tend to have wider spreads. They use this conclusion to suggest that the spread is positively related to information asymmetry. In addition to the cross sectional analysis, in this study we also link the spread and information asymmetry in a dynamic context. Namely, we show how the change of spread is related to the change of information asymmetry after controlling for other determinants of the spread such as the price of the security, trading volume, and return variance.

Thirdly, Dupont (2000) and Lee et al (1993) argued that the spread and the depth are the two main indicators of market liquidity. The spread represents the price dimension, while the depth represents the quantity dimension of liquidity. However, most previous studies related to the quoting behaviour of market makers neglected the importance of the depth. This study fills this gap. Particularly, we examine the intraday pattern of the quote depth and also test how the depth is affected by a change in information asymmetry in a dynamic context, which is not done by previous studies. Providing information on how market liquidity (including both spread and depth) changes within a trading day is useful to investors, academics and policy makers.

Chapter 3: Theoretical Issues and Hypotheses

In this study, we consider two groups of issues. First, we consider how information asymmetry, the spread, and the depth change during a trading day. For this purpose, we divide each trading day into 13 half-hour time intervals, measure information asymmetry, the spread, and the depth during each time interval, and compare levels of each variable across the 13 trading periods. Second, we examine how the spread and quote depth are related to the degree of information asymmetry on both an intraday dimension and a cross-sectional dimension.

Several studies argue theoretically that information asymmetry should be higher at the opening of trading and drop as trading goes on. Foster and Viswanathan (1990, p594) state: "because the price is an important source of information for uninformed liquidity traders, the informed trader has the greatest advantage when the market first opens." As trading goes on, through public information announcements and the inferences from order flow by market makers, the adverse selection problem should become less severe. Madhavan (1992) also argues that information asymmetry gradually dissolves as trading goes on and therefore information asymmetry at the end of a trading day should be lowest. Foster and Viswanathan (1994) constructed a model with competitive better informed and less informed traders. In their model the better informed traders strategically chose to trade at early periods based on the common information known by them and by the less informed traders. Through the trading process, the common information has been gradually dissipated and then the better informed traders will start to trade based on their additional information. Overall, we believe that the information asymmetry faced by market makers should demonstrate a distinct intraday pattern, particularly decreasing over trading hours. However, as mentioned by Foster and Viswanathan (1994), the strategy of better informed traders could lead to an increase in the degree of information asymmetry at certain periods of a trading day. These arguments lead to the following two hypotheses:

Hypothesis 1.a: *The degree of information asymmetry decreases through the trading day.*

Hypothesis 1.b: *The rate of decrease in information asymmetry starts high during the first trading period and drops gradually as time goes by.*

Most previous studies provide theoretical arguments suggesting that the spread should change throughout a trading day following a U-shaped pattern. At the morning, the spread should start at its highest level and should drop as trading goes on until later in the afternoon when the spread starts increasing again. Brock and Kleidon (1992) argue that the spread tends to be wide during the opening and closing hours due to the less elastic demand of investors. Foster and Viswanathan (1994) and Madhavan (1992) both argued that, through the trading process, information asymmetry gradually resolves, and therefore the degree of information asymmetry declines during a trading day. If all other components of the spread are constant it follows that the spread should be widest at the opening and declines throughout the day. However, previous studies, such as Amihud and Mendelson (1982), suggest that as the trading day approaches the closing minutes, inventory control issues tend to dominate the attention of market makers. Hence, the U-shaped pattern of the spread that is observed by prior studies is consistent with higher information asymmetry costs at the beginning of the day and increased inventory holding costs or increased order processing costs in the form of monopoly profit due to inelastic demand at the end of the day.

In a more recent study Ascioglun et al (2006) showed that the spread starts high as demonstrated by previous studies but it levels off during the rest of the trading day and does not rise as closing approaches. They argued that the intraday pattern of the spread has changed due to the dramatic changes in the marketplace. While Ascioglun et al (2006) stop short of indicating specifically what market changes may have altered the daily pattern of the spread, other authors point out at least one possibility. Chakravarty et al (2005), Chakravarty et al (2004) and Zhao and Chung (2006), showed that the spread, the quote depth and the degree of informed trading have all changed after the implementation of decimal pricing (on January 29, 2001, the NYSE converted all trading to decimal pricing from the 1/16 minimum quote size). This raises the questions of whether the intraday pattern of the spread has stayed the same as suggested by earlier research or changed significantly as suggested by Ascioglun et al (2006). In this study we attempt to resolve these questions. Specifically, we examine the following hypothesis:

Hypothesis 1.c: *During a trading day, the spread changes following a U-shaped pattern – it starts high during the morning period of a trading day, drops gradually as time goes by to* *reach a minimum sometime after midday, and rises as the closing time approaches.*

As suggested by Lee et al (1993), both the spread and the depth are important dimensions of liquidity where the spread reflects the price dimension and the depth reflects the quantity dimension. Their empirical results showed a negative relationship between the spread and depth around earnings disclosure events and also on an intraday basis. Particularly, higher bid-ask spread and lower depth will be quoted around times of higher informed trading. Dupont (2000) supports this conclusion by arguing that the quote depth can be used by market makers to fend off the information risk that arises from increased informed trading. He showed that the depth quoted by market makers is more sensitive to changes in the degree of information asymmetry than the spread. Li et al (2005) tested the patterns of depth quoted exclusively from specialists and limit orders from the NYSE and found an inverted U-shaped pattern for both of them. Since these studies suggested that the spread and the depth are negatively related and both are tools used by market makers to fend off the information risk, the depth is expected to rise over trading hours with the decrease in information asymmetry. However, the inverted U-shaped pattern found by these studies implies that the drop of depth toward the end of the day might be driven by other factors that are affecting market makers' quoting behaviour, such as inventory and monopoly effects. Therefore, we test the following hypothesis:

Hypothesis 1.d: *During a trading day, the depth changes following an inverted U-shaped pattern – it starts low during the morning period of a trading day, rises gradually as time goes by to reach a maximum sometime after midday, and drops as the closing time approaches.*

Previous researchers, such as Foster and Viswanathan (1990) and Glosten and Milgrom (1985) have suggested that the spread is partially determined by adverse selection costs. Therefore, when faced with a higher degree of information asymmetry, market makers will increase the spread to gain from liquidity traders in order to compensate the loss to informed traders. This reaction of market makers to the increased informed trading is not only on the price dimension but also on the quantity dimension. As Dupont (2000) and Li et al (2005) pointed out, the change of the spread is accompanied by the change in the quote depth as a tool for market makers to manage the information risk. Market makers can reduce the number of shares traded at a loss to informed traders by quoting lower depth when they find a higher degree of information asymmetry in the market. Therefore, the spread and the quote depth are expected to be respectively positively and negatively related to the degree of information asymmetry. Such relationship should remain the same in both intraday and cross sectional variation aspects.

We test the following hypotheses regarding the relationship between the spread and quote depth and the degree of information asymmetry.[2]

Hypothesis 2.a: *The spread quoted by market makers changes positively with the variation in information asymmetry.*

Hypothesis 2.b: *The depth quoted by market makers changes negatively with the variation in information asymmetry.*

[2] It is possible that information asymmetry may affect the market maker's optimal inventory level, hence the inventory holding costs. Our analysis accounts for these effects indirectly as we measure the overall impact of information asymmetry on the spread and the depth.

Chapter 4: Data

4.1 Data Sources and Sample Selection

This study considers the patterns of intraday movements of information asymmetry, the spread, and the depth for stocks listed on the NYSE. Prior studies have shown that due to different trading mechanisms and institutional characteristics, the spreads of stocks traded on the NYSE follow an intraday pattern different from the pattern followed by the spreads of stocks listed on the NASDAQ. We choose the NYSE listed stocks to be the target of our research. The NYSE listed stocks are more representative of the population of stocks given that the NASDAQ is typically known as a high-tech market, attracting many of the firms dealing with internet or electronics products.

We obtain the data for our study from the Trade and Quote (TAQ) database provided by the NYSE, which contains transaction-by-transaction data including records for both trades and quotes for all securities listed on the NYSE, AMEX, and NASDAQ exchanges. Our data sample begins January 2005 and ends December 2005. Each trade or quote record is time-stamped to the nearest second, which is ideal for the intraday study.

To construct our data sample, first we get from the CRSP database the list of securities trading on the NYSE during 2005. Initially the list included 2,668 securities. However, we exclude any issues that do not have a continuous or exclusive trading history on NYSE. This filter cuts the sample size to 1,725 securities. We exclude securities that are traded both on NYSE and any other exchange, because the level of information asymmetry could be affected by cross listing.[3] McInish and Wood (1992), Noronha et al (1996), and Tannous and Zhang (2008) argued that the cross listing of stocks could attract more informed traders since the opportunities of trading on private information are increased. Among the 1,725 securities, we only include common share issues. We feel that the information environment for common shares is different from the information environment of other types of issues such as preferred shares or American Depository Receipts. Furthermore, based on the Distribution Event Array

[3] It would be interesting as well to examine information asymmetry patterns for stocks listed on the NASDAQ and cross-listed stocks. We leave this task to future studies.

obtained from CRSP,[4] we also exclude stocks that undertook a stock split during 2005,[5] and stocks with closing prices less than $5 or greater than $150.[6] These filters reduce our data sample to 686 stocks. Among these 686 stocks, only 678 stocks have trade and quote data available on the TAQ database. Thus our data sample contains these 678 stocks. In order to estimate Equation (1) below, we require a minimum of ten observations per stock for each interval. This constraint reduced the sample size to 677 stocks.[7]

Based on the final stock list, we extract transaction-by-transaction data from the TAQ database for both the trade data file and the quote data file. The trade file includes data for the stock symbol, trading date, trading exchange, trading time, transaction price, trading size and a code for the trading condition. The quote file includes data for the stock symbol, trading date, trading exchange, trading time, quoted bid price, quoted bid size, quoted ask price, quoted ask size and a code for the quoting condition. All the trading and quoting data are stamped to the nearest second.

There are several steps needed to ensure that the quote and trade data stem from normal market conditions: Firstly, we retain only trade records with a trade indicator less than 3, which indicates a good trade. (We exclude trades coded out of time sequence and coded as involving an error or a correction.) We also exclude all trades that happened outside the normal trading hours of the NYSE (from 9:30 am to 4:00 pm). Secondly, for quote data, we only include the quotes that are eligible for inclusion in the National and National Association of Securities Dealers Best Bid and Offer Calculation (BBO-eligible).[8] Thirdly, following Chung

---

[4] The Distribution Event Array is a list of events describing cash dividends, capital adjustments, and other distributions made to shareholders.

[5] Van Ness et al (2001, *p*4) exclude all stocks that undertook a stock split. They justify this by stating "several researchers show that market microstructure properties change around stock splits, for example: Angel (1997), Conroy, Harris and Benet (1990), Lipson (2001), Schultz (2000)."

[6] This filter just excluded stocks with closing prices less than $5 or greater than $150. The intraday prices of stocks extracted from TAQ database could (and do) go below $5 or above $150.

[7] This is because the GMM estimation we apply to Equation (1) requires that there be at least ten observations for each regression.

[8] Such quotes (mode = 1, 2, 6, 10, or 12) are called BBO-eligible, which indicates active trading sessions. Quotes with other condition codes are BBO-ineligible, falling into 4 categories: closing quotation, trading halts quotation, pre-opening quotation, and non-firm quotation.

and Zhao (2003), we apply the following filters to reduce data entry errors: (1) we exclude trades if the trade price or trade size is less than or equal to zero, (2) we exclude trades if $|(p_t - p_{t-1})/p_{t-1}| > 0.1$, (3) we exclude quotes if either the bid or ask quote is less than or equal to zero, (4) we exclude quotes if either $|(b_t - b_{t-1})/b_{t-1}| > 0.1$ or $|(a_t - a_{t-1})/a_{t-1}| > 0.1$ where b_t and a_t indicate respectively bid and ask quotes, (5) we exclude quotes if either the bid depth or ask depth is less than or equal to zero, and (6) we exclude quotes if the spread is greater than $5 or less than $0.

4.2 Derivation of the Trade Indicator

To estimate the level of information asymmetry directly, we need to know for each trade whether it is initiated by a buyer or a seller. As the TAQ database does not contain direct information about the direction of a trade, we use the procedure proposed by Lee and Ready (1991) to derive whether a trade is buyer initiated or seller initiated. This procedure requires matching each trading transaction with the prevailing quote. Hasbrouck et al (1993) pointed out that the different reporting mechanisms of trades and quotes will result in an incorrect time sequence of the transaction data, and usually the quotes are recorded before the trades that initiated their revision. Therefore, prevailing quotes derived from the original sequence of time might be biased, leading to the misclassification of buy and sell orders. Following Lee and Ready (1991), for each trade we identify the prevailing quotes to be those most recent but at least five seconds old before each trade. Then, the procedure below recommended by Lee and Ready (1991) is applied to derive the trade indicator.

First, we calculate the prevailing quote midpoint.[9] A transaction is classified as a buy if the price is above the prevailing quote midpoint and a sell if the price is below the prevailing quote midpoint.

For trades where the transaction price is at the midpoint, the tick test is applied for the classification, where we infer the transaction direction by comparing the trade price with the preceding trade's price. Trades are classified into 4 categories based on the tick test: uptick, downtick, zero-uptick and zero-downtick. If a trade price is higher (lower) than the previous

[9] Midpoint = (bid + ask)/2. Bid and ask quotes here are the prevailing quotes corrected by the 5-second rule.

one, then it is identified as an uptick (downtick). If the trade price is the same as the previous one and the closest identified trade is an uptick (downtick), a trade is identified as a zero-uptick (zero-downtick). Finally, a trade is classified as a buy if it is an uptick or a zero-uptick and a trade is classified as a sell if it is a downtick or a zero-downtick.[10] A trade indicator is equal to +1 for a buyer-initiated trade and –1 for a seller-initiated trade.

After the data filter procedure and matching each transaction with the prevailing quote, our data sample contains 226,509,425 transaction level records, including stock symbol, trading date, trading time, trading price, trading size, prevailing bid price, prevailing ask price, prevailing bid size, prevailing ask size, and trade indicator.

[10] Through the tick test process, theoretically, all the trades within a day after the first identified trade of each day can be identified as buyer initiated or seller initiated.

Chapter 5: Intraday Pattern of Information Asymmetry

This chapter measures the degree of information asymmetry for each stock and each trading interval. Then, we use this information to determine and analyze the pattern of information asymmetry as trading progresses through a trading day.

5.1 Methodology

5.1.1 The Model for Measuring the Degree of Information Asymmetry

The model we use to estimate the degree of information asymmetry directly (as opposed to through a decomposition of the spread) is adapted from Madhavan and Smidt (1991) and Noronha et al (1996). It assumes the market maker uses Bayesian rules to update her beliefs about the future stock price. Thus, the expected value of the stock is modeled as a combination of the prior mean based on public information and a noisy signal based on the private information contained in the current order flow. That is to say, a price change might occur either in the case of a public announcement without any trade taking place or in the case of responding to the process of trading itself without any public announcement. We illustrate below how the following model supplies us with a direct way to measure the level of information asymmetry on an intraday basis:

$$\Delta p_{k,i,t} = \beta_0 + \beta_1 q_{k,i,t} + \beta_2 D_{k,i,t} + \beta_3 D_{k,i,t-1} + \varepsilon_{k,i,t}, \tag{1}$$

where $\Delta p_{k,i,t} = p_{k,i,t} - p_{k,i,t-1}$ represents the change of transaction price from one trade to the next, k indicates each of the 678 stocks in the sample, i indicates each of the 13 trading intervals, t indicates the order of trades in each interval, $q_{k,i,t}$ represents the signed transaction size,[11] $D_{k,i,t}$ is the trade direction indicator that takes the values +1 for a buy and –1 for a sell, and $\varepsilon_{k,i,t}$ is a stochastic error term.

In Equation (1) the coefficient β_1 captures the information effect, the responsiveness of price to order quantity. However it also includes the effect of costs varying with order size. Easley and O'Hara (1987), Glosten (1989) and Lin et al (1995) argued that market makers detect a higher degree of information asymmetry from large orders. Therefore, the coefficient β_1 is expected to be positive. Madhavan and Smidt (1991) proved that an alternative measure of the significance of information asymmetry is the weight placed on prior beliefs measured

[11] Signed transaction size is the trading shares for each transaction multiplied by the trade direction indicator.

by the PRIOR = $-\beta_3/\beta_2$ which can be estimated by model (1). The PRIOR actually represents the weight placed on public information. If order flow is uninformative, the weight placed on public information should be close to unity. However if there exists severe information asymmetry in the market, market makers will be more sensitive to the order flow, and the weight placed on public information approaches zero. This means that higher (lower) estimated PRIOR indicates lower (higher) degree of information asymmetry. The coefficients β_2 and β_3 are expected to be respectively positive and negative and should satisfy the size restriction: $\beta_2 > |\beta_3|$. Therefore, the PRIOR should be a positive number between zero and one. We use the index IA = 1 – PRIOR to represent the degree of information asymmetry.

In order to test the intraday pattern of the degree of information asymmetry, spread and quote depth, following Heflin et al (2007) and McInish and Wood (1992), we divide a trading day into 13 trading intervals to exhibit the intraday variation. The normal trading hours on the NYSE are from 9:30 am to 4:00 pm, which give 13 successive 30-minute intervals: 9:30am–10:00am, 10:00 am–10:30am, …, and 3:30pm–4:00pm, labelled intervals 1, 2, …, 13 respectively. Each transaction falls into one interval (whereas the corresponding quotes may fall into a previous interval).

5.1.2 Description of the Data

Table 2 provides descriptive statistics related to our sample of 677 NYSE stocks. It provides information on share price (transaction price per trade), share volume (number of shares traded), trade size (number of shares traded at a single transaction), and number of trades per interval. The share price, share volume and trade size are first computed for individual stocks across all transactions during trading days in 2005 and then the descriptive statistics are calculated from the stock-specific data. The number of trades per interval is first computed for each trading interval for each stock, and then the descriptive statistics are computed across the 13 trading intervals of the 677 stocks.

====== Insert Table 2 Here======

From Table 2, we see the average share price ranged from \$4.98 to \$153.59 with the mean value equal to approximately \$36 and the total share volume for the whole 2005 year ranges from 0.102 million shares to 4,497 million shares with average share volume equal to

240 million shares. For the average trade size, the smallest size is 145 shares per trade, the highest size is 3,671 shares per trade, with a mean of 539 shares per trade, and the number of trades per interval has a minimum of 16 trades for one interval, the maximum being 157,528 trades, and the mean is 25,737 trades.

5.2 Empirical Results for the Intraday Pattern of Information Asymmetry

5.2.1 Estimation of the Degree of Information Asymmetry

We have 677 NYSE stocks in our data sample, which means within each trading interval, we will have 677 groups of transaction-by-transaction data. Model (1) will be applied to each trading interval for each individual stock, and therefore we are going to have 677 × 13 = 8,801 equations to be estimated in total. First OLS will be applied to all the 8,801 equations.

Since we cannot provide the regression results for all 8,801 equations here, Panel A of Table 3 shows the frequency of positive, negative, significant or insignificant coefficients of all the independent variables in Equation (1) estimated by OLS. We can see that 8,707 estimates of parameter β_1 have the expected positive sign, 8,791 estimates of parameter β_2 have the expected positive sign, and 8,784 estimates of parameter β_3 have the expected negative sign. Moreover, the estimates of parameters β_1, β_2 and β_3 that are of the expected sign are mostly statistically significant at the 5% level or better (respectively, 92.55%, 99.32% and 98.58% are significant at the 1% level and 94.78%, 99.55% and 99.01% are significant at the 5% level). It also shows that 8,771 equations satisfy the size restriction $\beta_2 > |\beta_3|$. Thus, the vast majority of the estimates have the expected sign for all the coefficients and most of those are significant.

====== Insert Table 3 here======

According to previous empirical studies, for example Ascioglu et al (2006) and Heflin et al (2007), it is common for high frequency data to exhibit both heteroskedasticity and autocorrelation in the disturbance term, which can cause hypothesis testing based on statistics estimated by OLS to be unreliable. Since there is no appropriate way for us to identify the form of the heteroskedasticity and autocorrelation for all the 8,801 equations, following prior research (Ascioglu et al, 2006 and Heflin et al, 2007), we employ Hansen's (1982) generalized methods of moments (GMM) procedure to obtain consistent estimates of the covariance matrix.[12]

The regression results of Equation (1) based on the GMM estimation are reported in Panel B of Table 3, which shows the frequency of positive, negative, significant or insignificant coefficients of all the independent variables in Equation (1). Since we didn't use any instrumental variables[13] and we used the same independent variables for the GMM estimation as for OLS, the parameter estimates are the same; however, the t-statistics are generally different for these two estimation methods. At the 1%, 5%, and 10% significance levels, 67.08%, 99.26%, and 98.52%; 81.01%, 99.56%, and 98.95%; and 86.91%, 99.59%, and 99.16% of the estimates of parameters β_1, β_2, and β_3 are of the correct sign and significant respectively. The main effect from using GMM estimation is reducing the incidence of significance of the estimates for β_1. The incidences of significance of β_2 and β_3 are only slightly lower than they are with the OLS regression, so we have some confidence that the estimates of the PRIOR are not strongly affected by heteroskedasticity and autocorrelation.

5.2.2 The Intraday Pattern of Information Asymmetry and Analysis of Results

Based on the estimated values of β_2 and β_3 from the 8,801 equations, we can calculate the PRIOR = $-\beta_3/\beta_2$, which measures the weight placed by market makers on public information, to show the level of information asymmetry.[14] As indicated earlier, our regression is applied

[12] SAS supports GMM estimation in the proc model. In order to implement the GMM estimation, we need to specify the type of weighting matrix we are going to use. According to SAS/ETS User's Guide (*p*1011), we need to specify the kernel and the bandwidth to run the GMM estimation. There are three kinds of kernels supported by the proc model: the Bartlett Kernel, the Parzen Kernel and the Quadratic Spectral Kernel. In our estimation, we implement the Bartlett Kernel. The SAS document indicates that the Newey-West consistent covariance corresponds to the Bartlett kernel with bandwidth parameter $l(n) = L + 1$. The Newey-West covariance estimator is L then the corresponding bandwidth for the Bartlett kernel is $L + 1$. This provides us with an appropriate way to identify the bandwidth in the GMM estimation procedure. It is also pointed out that the bias of the standard error estimates increases for large bandwidth parameters. Therefore, to balance these we use a bandwidth of 10 to implement the GMM estimation. Other bandwidths are tested in Section 8.1.

[13] Since we are not using any instrumental variables for the GMM estimation, we do not specify the moment conditions here, following Heflin et al (2007).

[14] A small portion (42 out of 8,801) of the estimated parameters β_2 and β_3 either fail to have the expected sign or fail to meet the size restriction. Initially we do not exclude these regression results while calculating the PRIOR values for each stock within each trading interval. In Section 8.3 we report a robustness test by excluding these observations. This creates no material difference in the analysis.

to each of the 677 stocks in our data sample within each of the 13 trading intervals. Thus the PRIOR is calculated for each stock during each trading interval. This means, for each interval, we have 677 calculations of the PRIOR. Since the PRIOR represents the weight placed on public information, higher (lower) value of the PRIOR indicates lower (higher) degree of information asymmetry. We use the variable IA = 1 – PRIOR to measure the level of information asymmetry directly. The average process might be affected by the outliers, therefore besides the mean value of IA for each trading interval we also report the lower bound and the upper bound of the 95% confidence interval for the mean in Panel A of Table 4.

====== Insert Table 4 Here======

We observe that the mean value of IA declines gradually until interval 7 (12:30 noon to 1:00 pm), rises over trading intervals 8, 9, and 10 and then drops again monotonically during the remainder of the trading day. This indicates that the degree of information asymmetry is decreasing as trading hours goes by except for a small rise after interval 7. Information asymmetry is highest during the opening interval and lowest during the closing interval. Furthermore, according to the reported lower and upper bounds, we see that the confidence interval varies consistently with the variation of the mean value of IA. Figure 1a illustrates the intraday variation of information asymmetry based on the mean value of IA from Panel A in Table 4, including the lower and upper bounds of the 95% confidence interval.

======Insert Figure 1a Here ======

Roughly, the intraday pattern of information asymmetry shown in Figure 1a is in line with previous literature. For example, Madhavan (1992) pointed out that information is gradually dissolved as trading goes on and therefore information asymmetry at the beginning of a trading day should be highest and at the end of a trading day should be the lowest. One exception here is that after interval 7, the degree of information asymmetry rises through intervals 8 to 10. This observation might be due to less information being resolved through the trading process due to the lower trading volume during the noon period, and therefore information accumulates during this time. We plot the intraday variation of trading volume as Figure 1b.

======Insert Figure1b Here ======

Figure 1b shows that the trading volume is at higher level at the opening and at the closing and is at the lowest level at interval 7 (12:30 pm to 1:00 pm). Trading Interval 7 is the same interval where information asymmetry experiences a local minimum. The reported 95% confidence interval varies consistently with the mean value of trading volume. In addition, notes to Figure 1b report the t-test statistics which indicate that the trading volume during interval 7 is significantly lower than the trading volume during the period before and the period after. We speculate that a new wave of information may be arriving to the market after the noon hour following the conclusions of early morning decision meetings conducted by investment management committees, and other decision makers of companies. Another possible explanation for the increase in information asymmetry after the noon period could be the competition between better informed and less informed traders. Foster and Viswanathan (1994) argue that the better informed traders strategically chose to trade at early periods based on the common information known both by them and by the less informed traders. When the common information has been resolved through trading, the better informed traders will start to trade based on additional information known only by them. Such a trading strategy by better informed traders could cause the accumulation of information at certain times within a trading day.

In order to test Hypotheses 1a and 1b, we use dummy variables to represent the different trading intervals and run the following regression to compare the degree of information asymmetry across the 13 trading intervals. Here, we choose interval 7 as the benchmark. Equation (2) is used for this analysis:

$$IA_{k,i} = \gamma_0 + \sum_{i=1}^{6} \gamma_i DUM_i + \sum_{i=8}^{13} \gamma_i DUM_i + \varepsilon_{k,i}, \tag{2}$$

where $IA_{k,i}$ is the measure of information asymmetry for each stock k within each interval i calculated from the regression results of Equation (1) and DUM_i is a dummy variable that represents the trading intervals 1 to 6 and 8 to 13 taking the value one if an observation occurs during interval i and zero otherwise. The intercept captures interval 7, 12:30 am to 1:00 pm.

Panel B of Table 4 reports the regression results. We see that the degree of information asymmetry for earlier trading intervals and later trading intervals are respectively significantly higher than and lower than trading interval 7. In particular, information asymmetry during trading intervals 1 to 6 is significantly higher than information asymmetry during trading interval 7. In contrast, information asymmetry during trading interval 13 is significantly lower than information asymmetry during trading interval 7, which is consistent with Hypothesis 1a. This result implies that market makers face the greatest informational disadvantage at the beginning of a trading day and the least informational disadvantage at the end of the trading day. Particularly, the drops of information asymmetry from interval 1 to interval 2 and from interval 12 to interval 13 are the greatest according to the estimated parameters of the dummy variables, which is consistent with Figure 1a. However, we also find that the difference between information asymmetry during interval 10 and information asymmetry during interval 7 is significant at the 1% level. This observation indicates that information asymmetry does indeed increase after the noon hour.

Chapter 6: Intraday Patterns of Spread and Depth

Previous studies, such as Dupont (2000), Gong (2007), Gregoriou et al (2005), Lee ea al (1993) and McInish and Wood (1992) argue that in response to an increase in the degree of information asymmetry, market makers may widen the spread or lower the depth or both to mitigate the loss to traders with superior information. This chapter examines how the spread and quote depth vary over the trading hours within a day and compares the results to the intraday variations in information asymmetry.

6.1 Intraday Pattern of Spread

Consistent with the intraday tests for information asymmetry, we analyse the spread over 13 trading intervals in each day. Within each trading interval, we consider the intraday data of the spread and the quote depth for the sample of 677 stocks. First, we focus on the time-weighted average relative spread. It is calculated as follows:

$$SPRD_{k,i} = \sum_{q=1}^{Q} w_{k,i,q} \times \frac{askprice_{k,i,q} - bidprice_{k,i,q}}{0.5 \times (askprice_{k,i,q} + bidprice_{k,i,q})}, \tag{3}$$

where $SPRD_{k,i}$ is the time-weighted average relative spread for stock k during trading interval i, $askprice_{k,i,q}$ and $bidprice_{k,i,q}$ are respectively the quoted ask and bid price of quotation q of interval i for stock k, Q is the number of quotations for security k during interval i, and $w_{k,i,q}$ is the time weight calculated as the number of seconds quote q is outstanding divided by the total number of seconds of interval i.[15] Panel A of Table 5 shows the descriptive statistics of SPRD for the 677 stocks across the 13 trading intervals.

======Insert Table 5 Here======

From the column of mean value, we see that SPRD is at the highest level at the opening interval and then it decreases throughout the trading day, with the exception of small increases for the 9th and 13th intervals. It reaches the minimum value at trading interval 12. As with IA, we also report the 95% confidence interval for the mean values of SPRD. The lower and upper bounds of the confidence interval for the mean of SPRD follow the same pattern as the mean. Figure 2 illustrates these observations.

[15] All the calculations for Equation (3) are based on the dataset with the trade data and prevailing quote data matched using the five-second rule.

====== Insert Figure 2 Here======

The same empirical test of intraday variation of AI is also applied to SPRD. The equation is as follows:

$$SPRD_{k,i} = \delta_0 + \sum_{i=1}^{6} \delta_i DUM_i + \sum_{i=8}^{13} \delta_i DUM_i + \varepsilon_{k,i}, \tag{4}$$

where $SPRD_{k,i}$ is as defined in Equation (3) and the dummy variables are as defined in Equation (2). The regression results are reported in Panel B of Table 5. Compared with interval 7, which is captured by the intercept, SPRD is significantly higher during trading intervals 1 and 2 at the 1% level, (and interval 3 at the 10% level), but it is not significantly different for the remaining intervals. Our finding of the intraday pattern of spread is consistent with Ascioglun et al's (2006) research. They showed that the spread drops sharply after the early two trading intervals and then remains fairly stable during the rest of the trading day. Moreover, they attribute the change in the spread towards the end of a trading day to changes in the marketplace. Note that this is quite different from the intraday pattern of information asymmetry shown above, where the degree of information asymmetry drops sharply during the last trading interval. The decoupling of SPRD and IA, suggests that other costs of market making, such as inventory holding costs and the monopoly power of the market maker suggested by Amihud and Mendelson (1982) and Block and Kleidon (1992), may have more significant roles towards the end of the trading day. For example, the inventory holding costs include the possibility of capital losses should the prices of the securities held by the market maker drop significantly in the future. This risk becomes higher during the last trading interval of the day given the long period of non-trading hours that follows. Consequently, the market maker seems to keep the spread during the last trading period constant despite the drop in the adverse selection costs. The additional profits from lower adverse selection costs during this period compensate for the higher inventory risk.

However, according to the intraday pattern of depth we derive later, it seems that the inventory holding costs are a main consideration since the depth rises to its highest level during the closing interval. Another possible explanation is that during the last trading interval the rate of arrival of limit orders slows down which leaves the market maker greater monopoly power to keep the spread constant despite the reduction in the adverse selection costs. As is shown by Chung et al (1999), compared with the midday, there are less limit orders outstanding at the end of a trading day.

6.2 Intraday Pattern of Depth

The same analysis performed on SPRD is also applied to the quote depth. The time-weighted average depth for each stock during each trading interval is computed as follows:[16]

$$DPTH_{k,i} = \sum_{q=1}^{Q} w_{k,i,q} \times \frac{biddepth_{k,i,q} + askdepth_{k,i,q}}{2}, \tag{5}$$

where $DPTH_{k,i}$ represents the time-weighted average depth for stock k during the trading interval i, $biddepth_{k,i,q}$ ($askdepth_{k,i,q}$) is the available number of shares the market maker is willing to buy (sell) at the bid (ask) price of quotation q of interval i for stock k, and w is as defined in Equation (3).

Panel A of Table 6 reports the descriptive statistics of DPTH for the 677 stocks across the 13 trading intervals. The mean value, the lower bound value and the upper bound value for each interval are computed. In contrast with the trend of SPRD, the mean value of DPTH increases throughout the trading day. The reported lower and upper bounds of a 95% confidence interval behave similarly. This is illustrated in Figure 3.

======Insert Table 6 and Figure 3 Here======

As with IA and SPRD, the test for differences in the intraday pattern is applied to DPTH. The equation is as follows:

$$DPTH_{k,i} = \theta_0 + \sum_{i=1}^{6} \theta_i DUM_i + \sum_{i=8}^{13} \theta_i DUM_i + \varepsilon_{k,i}\,. \tag{6}$$

The estimated results are shown in Panel B of Table 6. We find that DPTH increases monotonically throughout the trading day. During the first trading interval it is at its lowest level, and it is significantly lower than DPTH during Interval 7. In contrast, during Intervals 12 and

[16] The measure of depth in units of the number of shares might be sensitive to the price level (see Table 14 below). Another useful measure could be the dollar value of stock that can be traded on each side (ask depth times ask price and bid depth times bid price). However, we use the number of shares as a measure of the depth to be consistent with previous studies such as Lee and Ready (1993) and Tannous and Zhang (2008).

13, DPTH is significantly higher than that during Interval 7, suggesting that DPTH reaches its highest level at the end of a trading day. However, during Intervals 2-6 and 8-11 DPTH is not significantly different from that during Interval 7. Overall, DPTH rises significantly from the beginning of a trading day toward the end of a trading day, which is different from the inverted U-shaped pattern reported by previous studies such as Li et al (2005). DPTH does not drop at the closing interval. Instead, it rises significantly. Here, the behaviour of market makers on the depth dimension is consistent with the intraday variation of information asymmetry. As we have mentioned earlier, while market makers detect higher (lower) informed traders, they are going to decrease (increase) the number of shares they are willing to trade at the quoted price. Therefore, the sharp increase of DPTH at the closing trading interval might be explained by the sharp drop in the degree of information asymmetry at the end of the trading day. In addition, it is consistent with the inventory rebalancing and risk reduction motives of the market maker. As closing approaches, it is logical for the market marker to reduce inventories to avoid the risk of holding a large position overnight during which price may decrease significantly. Therefore, in response to the drop in information asymmetry near closing the market maker would increase the depth to reduce inventories before closing or if the inventory is already at optimal level would simply quote from the books to avoid increasing the inventory. Note that this behaviour is possible due to the monopoly power of the market maker. Without such power the market maker would have to reduce the spread to keep the marginal revenue equal to the marginal cost.

Chapter 7: The Relationships between Information Asymmetry, Spread, and Depth

Previous studies, for example McInish and Wood (1992) and Tannous and Zhang (2008), have shown that trading volume, return variance, and price level significantly explain the cross sectional variations of the spread. First, they argue that higher trading volume will make it easier for market makers to rebalance their inventories and therefore the reduced inventory cost can lead to a lower spread. Thus, the spread is expected to be negatively related to trading volume. Second, for stocks with higher return variance, market makers face higher inventory holding risk, which can lead to a higher spread. Therefore, we expect a positive relationship between the spread and the variance level. Third, for stocks with higher price level the fixed processing costs are spread over a higher value. Thus, we expect the relation between the spread and the price level to be negative. In another vein, previous studies suggest that a significant determinant of the spread for a security is the degree of information asymmetry this security's market maker faces. This chapter extends this research by examining the relationship between the spread and the degree of information asymmetry in the presence of controls for the trading volume, return variance and price level. We test this relationship separately across the trading intervals within a day and across the stocks. Given the significant intraday variations in information asymmetry we feel that it is important to extend our analysis to the intraday relationship. We perform similar analysis for the depth. Previous studies (Dupont, 2000 and Lee et al, 1993) have found that a higher spread is accompanied by a lower quote depth while the degree of information asymmetry is higher. However, the positive relationship between the spread and the degree of information asymmetry and the negative relationship between the quote depth and the degree of information asymmetry are not tested in a dynamic context. By testing how the intraday variations of the spread and the quote depth are related to the intraday variation of information asymmetry, we exhibit the dynamic effect of information asymmetry on the market making strategy of market makers.

7.1 The Relationship between Spread and Information Asymmetry

7.1.1 Intraday Variation Relationship

For the spread, the test model is as follows:

$$SPRD_{k,i} = \mu_0 + \mu_1 IA_{k,i} + \mu_2 VOL_{k,i} + \mu_3 VAR_{k,i} + \mu_4 PRC_{k,i} + \varepsilon_{k,i}, \tag{7}$$

where $SPRD_{k,i}$ is the time-weighted average relative spread defined in Equation (3), $IA_{k,i}$ is the measure of information asymmetry defined in Equation (2), $VOL_{k,i}$ is the total trading volume measured as the number of trading shares of stock *k* in trading interval *i*, $VAR_{k,i}$ is the return variance calculated from the transaction-by-transaction return for each stock *k* within trading each trading interval *i*, $PRC_{k,i}$ is the average transaction price per share of stock *k* during trading interval *i*, and $\varepsilon_{k,i}$ is the stochastic error term.

In Equation (7), the trading volume, return variance, and share price are included as control variables and our main interest is in the parameter μ_1. Since all the variables in Equation (7) are calculated for each stock within each trading interval, our data should be viewed as a panel structure: the data for each stock within each interval contributes to each observation.[17] The data at the stock level produce the cross sectional effect and the data through the intervals produce the time series effect. We estimate Equation (7) to capture the intraday variation effect and the cross sectional variation effect. Here we test the intraday variation of SPRD by controlling the effect of variation among different stocks.

For panel data, the two main models that can be applied are the fixed effects model and the random effects model.[18] We use the fixed effects model to estimate Equation (7).[19] Dummy variables on the cross section are commonly used to control the group effects in the fixed effects model, when the number of cross sectional units is small enough. However, when the cross-sectional units are substantial (our panel data has 677 cross sectional units), a transformed form of the estimation equation is needed to avoid computational problems. The

[17] Kennedy (2003, *p*302) defined panel data "as cross-sectional data in which we have observations on the same units in several different time periods. Panel data create more variability, through combining variation across micro units with variation over time."

[18] The fixed effects model examines group differences in intercepts, assuming the same slopes and constant variance across groups. The random effects model, by contrast, estimates variance components for groups and the error term, assuming the same intercept and slopes. The difference among groups lies in the variance of the error term.

[19] Kennedy (2003, *p*306) writes a procedure that shows how to make an appropriate choice between the fixed effects and random effects models. For our panel data, the F-test for no fixed effect is rejected, and therefore we choose the fixed effects model to do the estimation.

transformed fixed effects model derived from Equation (7) is as follows:[20]

$$(SPRD_{k,i} - \overline{SPRD}_k) = \mu_1(IA_{k,i} - \overline{IA}_k) + \mu_2(VOL_{k,i} - \overline{VOL}_k) + \mu_3(VAR_{k,i} - \overline{VAR}_k) + \mu_4(PRC_{k,i} - \overline{PRC}_k) + \varepsilon_{k,i}\,. \tag{8}$$

For each variable in Equation (7), the transformation is done by subtracting from each observation the average value of that variable within the cross sectional unit (here the cross sectional units correspond to the 677 stocks and each unit has 13 observations corresponding to the 13 trading intervals). For example, in Equation (8), each observation of each variable for stock k is reduced by the average value of that variable across the 13 trading intervals. Therefore, the new variables in Equation (8) are the deviations from group means. Using variables: $\widehat{SPRD}$, $\widehat{IA}$, $\widehat{VOL}$, $\widehat{VAR}$ and $\widehat{PRC}$ to represent the deviations from group means, the new equation is as follows:

$$\widehat{SPRD}_{k,i} = \mu_1\widehat{IA}_{k,i} + \mu_2\widehat{VOL}_{k,i} + \mu_3\widehat{VAR}_{k,i} + \mu_4\widehat{PRC}_{k,i} + \varepsilon_{k,i}. \tag{9}$$

According to Kennedy (2003) there are two kinds of variations in our panel data. One type is the variation in observations from individual unit to individual unit, and the other type is the variation from observation to observation within a single cross sectional unit. Equation (9) uses the first type of variation (the variation within each group), therefore sometimes the estimators are called the "within effects" estimators. Applied to our panel data, Equation (9) will test how the variation of spread from trading intervals 1 to 13 corresponds to the intraday variation of trading volume, return variance, price and the degree of information asymmetry across the 13 trading intervals, controlling the cross sectional variation by subtracting the group mean for each individual unit.

Table 7 reports the means for the spread, trading volume, return variance, share price and degree of information asymmetry for each of the 13 half-hour trading intervals. The intraday patterns of SPRD and IA have already been analyzed earlier. Trading volume is higher

[20] The estimation of parameters in Equation (9) will be the same as the estimation from inputting 677 dummy variables in Equation (7) representing the different unit effect (677 stocks), (see Greene, 2002). However, since no dummy is used, Equation (9) has more degrees of freedom for error, resulting in a smaller MSE (mean square error) and incorrect (larger) standard errors of parameter estimates. Thus we need to adjust the standard error by correcting the degree of freedom. Furthermore, the R^2 statistic for the within effects model is not correct because the intercept is suppressed. These corrections can be done by SAS in the TSCSREG procedure.

during the opening and closing trading intervals than it is at mid-day, and return variance is highest in the opening interval with little difference during the rest of the trading day. For the average price, little intraday variation is exhibited.

======Insert Table 7 Here======

OLS is applied to Equation (9) and Table 8 reports the results.[21]

======Insert Table 8 Here======

The estimated coefficient of the information asymmetry variable is significant and positive, which, indicates that when the degree of information asymmetry is higher, the spread is higher. This is consistent with previous research: when the degree of information asymmetry is higher, the adverse selection component of the spread should be larger. That leads to the wider spread. We find a significant positive relationship between the spread and the variable return variance, which is consistent with previous studies suggesting that a higher return volatility implies higher inventory risk and therefore results in a higher spread. The trading volume is found to be positively related to the spread, which is different from the findings of some previous cross sectional studies, for example McInish and Wood (1992) and Tannous and Zhang (2008), but is consistent with Lee et al's (1993) study for the intraday variation relationship between the spread and the trading volume.

Following Tannous and Zhang (2008), we also examined the correlations among the independent variables of Equation (9) to see whether there is a multicollinearity problem. Table 9 exhibits the Pearson correlation coefficient for each pair of the independent variables. We see that the variable IA is significantly correlated to the other independent variables. In particular, the information asymmetry variable is negatively related to the trading volume and price with correlation coefficients respectively equal to –0.04433 and –0.03529, both signifi-

---

[21] According to SAS/ETS User's Guide (*p*1728), the TSCSREG (Time Series Cross Section Regression) procedure to analyze a class of linear econometric models that commonly arise when time series and cross sectional data are combined. Therefore, we use the TSCSREG procedure to deal with the panel data we have here, consisting of time series observations on each of several cross-sectional units. One option of the TSCSREG procedure is FIXONE, which is employed to estimate Equation (9). However, the TSCSREG procedure with the FIXONE option actually reports results of least square regression with dummy variables, even though we don't need to create the dummy variables. Therefore the procedure reports the "correct" R^2 and standard errors, (the statistics are the same as those estimated by least squares with dummy variables).

cant at the 1% level. In contrast, it is positively related to return variance with the correlation coefficient equal to 0.02163, significant at the 5% level. However the correlations are generally quite low in absolute value ranging from –0.05 to 0.02, which suggests that these correlations are not economically strong.

======Insert Table 9 Here======

7.1.2 Cross Sectional Variation Relationship

In this part, we test whether the variable IA can also explain the cross sectional variation of the spread with trading volume, return variance and share price as control variables. This can be done by employing the fixed time effects model for our panel data, which will control the time effect for each trading interval.[22] This will be different from the model we used to test the intraday variation of spread, which uses dummy variables to control the differences among cross sectional units. Since there are only 13 trading intervals in our panel data sample, the time periods are small enough for us to employ dummy variables to represent each interval instead of using the transformed form as Equation (9). The test model for the cross sectional variation is as follows:

$$SPRD_{k,i} = \pi_0 + \pi_1 IA_{k,i} + \pi_2 VOL_{k,i} + \pi_3 VAR_{k,i} + \pi_4 PRC_{k,i} + \sum_{i=2}^{13} \alpha_i DUM_i + \varepsilon_{k,i}, \qquad (10)$$

where $SPRD_{k,i}$, $IA_{k,i}$, $VOL_{k,i}$, $VAR_{k,i}$ and $PRC_{k,i}$ are as defined in Equation (7), DUM_i is a dummy variable that takes the value 1 if the observation occurs during interval i and zero otherwise. The dummy variables are numbered 2–13, and the intercept will capture the effect of interval 1. The dummy variables in Equation (10) control the time effect of each interval on the spread so the parameters π_1–π_4 exhibit how the spread varies with the cross sectional variation of the variables information asymmetry, trading volume, return variance and share price.

Table 10 reports the cross sectional descriptive statistics for the variables in Equation (10), including the cross sectional variation for each variable. The spread ranges from a minimum value 0.0002578 to a maximum value 0.01485, the degree of information asymme-

[22] The fixed time effects model investigates how time affects the intercept using time dummy variables.

try ranges from 0.1161 to 0.7857, the trading volume varies from 8,707 shares to 345,974,892 shares, the return variance ranges from 0.00000001952 to 0.00003272, and the share price varies from 4.99 to 153.76. Note that Equation (10) is designed to examine how the spread is related to the independent variables after we control for the time effect. In particular, our main interest is in the parameter π_1, which displays the cross sectional relationship between the spread and the degree of information asymmetry. The relation has not been examined in this way by prior research. Previous cross sectional studies, for example Gregoriou et al (2005), did not directly measure the degree of information asymmetry, and instead used the level of disagreement in analysts' earnings forecasts as a proxy of information asymmetry.

======Insert Table 10 Here======

The regression results for Equation (10) are exhibited in Table 11. The estimated coefficient for the degree of information asymmetry shows a significant positive relationship between it and the spread. To illustrate, stocks that have high degree of information asymmetry tend to have a wider spread to make up the loss to the informed traders with superior information, which is consistent with Gong (2007) and Gregoriou et al (2005). Combined with the estimated results for Equation (9), the spread varies positively with both the intraday variation and cross sectional variation of the degree of information asymmetry, which is consistent with Hypothesis 2a. From the other estimated coefficients we can see that the trading volume and the share price are significant (at the 0.01 level) and negatively related to the spread. This is consistent with prior empirical results of McInish and Wood (1992) and Tannous and Zhang (2008). The variable return variance is found to be significant (at the 0.01 level) and positively related to the spread, which is consistent with the idea that market makers will widen the spread to manage inventory risk due to higher volatility.

======Insert Table 11 Here======

7.2 The Relationship between Depth and Information Asymmetry

According to Lee et al (1993), when market makers manage information risk, they cannot quote a very wide spread because they are employed to maintain market liquidity. Quoting a very wide spread might drive uninformed traders away and reduce market liquidity. Thus, the other alternative for market makers to protect themselves from information risk is to limit the number of shares they're willing to trade. We expect a negative relationship between the quote depth and the information asymmetry. We examine this hypothesis by applying to the depth the same analysis we applied to the spread. The trading volume, return variance and price level are also included as control variables. Higher trading volume implies higher liquidity as suggested by Harris and Raviv (1993) and therefore we expect the depth to be positively related to the trading volume. Higher return variance increases inventory holding costs, which should lead to a lower quote depth.

7.2.1 Intraday Variation Relationship

The model based on Equation (9) will change to:

$$\widehat{DPTH}_{k,i} = \rho_1 \widehat{IA}_{k,i} + \rho_2 \widehat{VOL}_{k,i} + \rho_3 \widehat{VAR}_{k,i} + \rho_4 \widehat{PRC}_{k,i} + \varepsilon_{k,i}, \qquad (11)$$

where $DPTH_{k,i}$ is the time-weighted average depth defined in Equation (5), $\widehat{DPTH}_{k,i}$ is the deviation of depth from its group mean, and the independent variables are the same as in Equation (9). Subtracting the group mean from each observation of each variable rules out the cross sectional variation effect. Thus, we can test the intraday pattern between quote depth and the degree of information, trading volume, return variance and price.

Table 12 shows the descriptive statistics for the depth, degree of information asymmetry, return variance, share price, and trading volume during each trading interval. DPTH is increasing consistently as trading hours go by. This pattern is opposite to the pattern of SPRD shown in Table 6. Next, we examine how the intraday variation of the degree of information asymmetry will affect the quote depth after controlling for trading volume, return variance and share price.

======Insert Table 12 Here======

The coefficients of Equation (11) are estimated using OLS and the results are reported in Table 13. Our primary interest is in the estimated coefficient of parameter ρ_1, which reveals the effect of the intraday variation of the degree of information asymmetry on the intraday pattern of the depth. As we can see in Table 13, ρ_1 is estimated to be negative and significant at the 0.01 level. Thus, during the opening interval of each trading day when the degree of information asymmetry is relatively high, the depth of the quotes is relatively low. When the degree of information asymmetry drops through trading, market makers are willing to provide more liquidity by increasing the depth. Thus, we find that the quote depth serves as a way for market makers to manage the risk of the variation of information asymmetry in a dynamic context. Quote depth is found to be positively related to the volume, which is consistent with our expectation. The coefficient of the return variance is estimated to be negative and significant at the 0.01 level. This implies that high variations negatively affect the quote depth. Higher return variance means higher inventory risk, and therefore market makers would prefer to be less involved in trading with investors when higher return variance is detected, which results in a lower posted depth from the market makers. Finally, the depth is shown to be negatively related to the share price suggesting that market makers are likely to offer greater depth for lower priced shares.

======Insert Table 13 Here======

7.2.2 Cross Sectional Variation Relationship

Furthermore, we also examine how the depth quoted by market makers varies from stock to stock with the cross sectional variation of the degree of information asymmetry, controlling the variations of trading volume, return variance and share price. Equation (10) will be applied to the dependent variable DPTH as follows:

$$DPTH_{k,i} = \vartheta_0 + \vartheta_1 IA_{k,i} + \vartheta_2 VOL_{k,i} + \vartheta_3 VAR_{k,i} + \vartheta_4 PRC_{k,i} + \sum_{i=2}^{13} \alpha_i DUM_i + \varepsilon_{k,i}, \qquad (12)$$

where $DPTH_{k,i}$ is the time-weighted average depth defined in Equation (5), and the other variables: $IA_{k,i}$, $VOL_{k,i}$, $VAR_{k,i,}$,$PRC_{k,i}$ and DUM_i are all defined in Equation (10). Here, the dummy variables control the intraday variation of quote depth. Parameters $\vartheta_1 - \vartheta_4$ capture how the cross sectional variation of the independent variables in Equation (12) affect the quote depth. OLS is used to determine the coefficients of Equation (12).

Table 14 demonstrates the regression results for Equation (12). Our primary interest is in the degree of information asymmetry. The estimated coefficient of IA is negative and significant at 1% level. It shows that the quote depth is negatively related to the degree of information asymmetry. It illustrates that the depth quoted by market makers will be lower for those stocks with a higher degree of information asymmetry. That is, market makers are willing to provide more liquidity for stocks with fewer informed traders. The coefficient of the variable trading volume is reported to be positive and significant. This implies that for stocks that trade actively the market maker is willing to provide greater depth. Perhaps it is easier for market makers to manage their inventories for highly active stocks. Furthermore, Table 14 shows that the quote depth is negatively related to the return variance, although the relationship is not significant. In contrast, the quote depth is shown to be significantly negatively related to the share price.

======Insert Table 14 Here======

7.3 Examination of the Relationship between Spread and Depth

Previous studies (Lee et al, 1993 and Li et al, 2005) suggest that a higher spread is accompanied by lower depth. From the empirical results above, we verify a greater spread and lower depth are associated with a higher degree of information asymmetry. That is to say, the spread and the depth are expected to have a negative relationship. To test this relationship, here we calculate and demonstrate the correlations between the spread and the depth for each of the 13 intervals in Table 15.

======Insert Table 15 Here======

Table 15 shows that for all the 13 intervals, the spread is found to be negatively related to the depth. However, the negative correlations are only significant during Intervals 1 to 5 and during Interval 13. For the remaining intervals, the negative correlations are found to be insignificant.

Chapter 8: Robustness Tests

8.1 Robustness Test 1: Different Bandwidths for GMM Estimation

While estimating Equation (1) by GMM, we have mentioned that we apply a bandwidth of 10 to each of our 8,801 regressions. However, different bandwidths can be calculated for each regression.[23] For robustness, we run the 8,801 regressions with bandwidths derived based on the number of observations for each equation. The regression results are reported in Table 16.

======Insert Table 16 Here======

Table 16 shows that the numbers and percentages of significant and insignificant estimates based on GMM estimations with different bandwidths are substantially similar to the results reported in Panel B of Table 3. Therefore, the conclusions we draw are robust to the choice of bandwidth used to estimate the IA variables.

8.2 Robustness Test 2: Subsamples Based on Trading Volume

Previous studies have shown that differences in trading activity may have an impact on the degree of information asymmetry, spread and quote depth, (see Ascioglu et al, 2006 and Lin et al, 1995). For securities that enjoy active trading, the price adjustment to new information is likely to be quicker than the adjustment for securities that are less actively traded. Thus, a higher trading volume should mean a lower degree of information asymmetry. Moreover, as we have discussed earlier, stocks with higher trading volume are quoted by market makers with a lower spread and a higher depth. In order to examine the magnitude of such effects, we divide our data sample into three groups according to the level of trading volume. We rank the 677 stocks according to their total trading volume throughout the whole year 2005, calculated based on the data extracted from TAQ, and group them as follows: The first 225 stocks will fall into the high activity group, the next 227 stocks will be in the medium activity group and the last 225 stocks will be in the low activity group.

In Table 17, the average value of the degree of information asymmetry, spread, quote depth are reported for each of the three groups. Our main purpose is to examine how the degree of information asymmetry, spread, and quote depth vary across the three groups. From

[23] SAS/ETS User's Guide (*p*1013) indicates that the bandwidth will be equal to $0.5\sqrt[3]{n}$ for the Bartlett Kernel where *n* refers to the number of observations for each equation.

the table, we can see that the level of information asymmetry is at the highest level in the low activity group, at the lowest level in the high activity group, and the medium level in the medium activity group. It is illustrated that for those thinly traded stocks that information asymmetry is high due to the lower level of trading volume, compared with those heavily traded stocks which have greater informational efficiency. This is consistent with the results obtained by Lin et al (1995). They suggest that the lower level of information asymmetry for high-volume firms could be the result of "a high level of monitoring and public information production by securities analysts and others" (p1166).

Consistent with the cross sectional variation test we have done, the spread is negatively related to the trading volume by groups: ranking from highest level, medium level to lowest level corresponding to the low activity group, medium activity group and high activity group. The average level of the depth varies positively with the level of trading volume by groups, which is also consistent with the cross sectional variation test. In addition, we conducted t-test to compare the means of the three different groups and the t-statistics are reported in Table 17. We can see that the means of IA, SPRD and DPTH for the three groups are all significantly different from each other at the 0.01 level with one exception. The t-test for the difference between the mean IA of the medium activity group and the mean IA of the low activity group is not significant.

======Insert Table 17 Here======

We also examine the intraday patterns of information asymmetry, spread, and quote depth when grouped by trading volume. From Figures 4a to 6c, the intraday patterns of information asymmetry, spread and quote depth for high activity, medium activity and low activity groups are demonstrated. As we can see, the intraday patterns for the high activity and medium activity groups basically remain the same as the intraday pattern based on the whole data sample. However, for the low activity group, the average changes of the degree of information asymmetry, spread, and quote depth are more volatile within a trading day. It illustrates that for stocks traded less frequently each trade tends to have greater price impact, which results in the high fluctuation of the pattern of the information asymmetry, spread, and quote depth.

======Insert Figures 4a to 6c======

Furthermore, we test whether the relationship between the spread and the information asymmetry and the relationship between the quote depth and the information asymmetry for the three groups of stocks remain consistent with the whole sample result. The test equations are as follows:

$$SPRD_H = \tau_0 + \tau_1 IA_H + \tau_2 VOL_H + \tau_3 VAR_H + \tau_4 PRC_H + \varepsilon_H \tag{13.1}$$

$$SPRD_M = \tau_0 + \tau_1 IA_M + \tau_2 VOL_M + \tau_3 VAR_M + \tau_4 PRC_M + \varepsilon_M \tag{13.2}$$

$$SPRD_L = \tau_0 + \tau_1 IA_L + \tau_2 VOL_L + \tau_3 VAR_L + \tau_4 PRC_L + \varepsilon_L \tag{13.3}$$

$$DPTH_H = \varphi_0 + \varphi_1 IA_H + \varphi_2 VOL_H + \varphi_3 VAR_H + \varphi_4 PRC_H + \varepsilon_H \tag{13.4}$$

$$DPTH_M = \varphi_0 + \varphi_1 IA_M + \varphi_2 VOL_M + \varphi_3 VAR_M + \varphi_4 PRC_M + \varepsilon_M \tag{13.5}$$

$$DPTH_L = \varphi_0 + \varphi_1 IA_L + \varphi_2 VOL_L + \varphi_3 VAR_L + \varphi_4 PRC_L + \varepsilon_L. \tag{13.6}$$

Equations (13.1) to (13.3) examine respectively how the spread is related to the information asymmetry for high, medium, and low activity groups of stocks while we control for the effects of price, volume, and return variance. Equations (13.4) to (13.6) are designed to conduct the same analysis for the depth. Table 18 shows the regression results for Equations (13.1) to (13.6). From the estimated coefficient of variable information asymmetry, we can see that for all the three groups, the spread and the quote depth are respectively positively and negatively related to the degree of information asymmetry. This observation is consistent with the analysis based on the whole sample.

======Insert Table 18 Here======

8.3 Robustness Test 3: Different Sample Size

While constructing our data sample, we exclude the stocks that have less than a minimum of 10 observations in each interval. This is done to satisfy the constraints of the GMM estimation we applied to Equation (1). For robustness, we check what happens to the results if we exclude those stocks with less than a minimum of 100 observations in each interval. The objective is to find whether the inclusion of inactive stocks biases our results regarding IA, since each estimate of IA contributes equally to the analysis, regardless of how precisely it is estimated.

In addition, when we analyze the regression results for Equation (1), we indicate that there are 42 equations out of the total 8,801 equations that have the wrong sign for parameter β_2 or parameter β_3 or do not satisfy the size restriction: $\beta_2 > |\beta_3|$, which leads to values of the PRIOR outside of the theoretical range of [0, 1]. Here also we exclude these 42 equations for robustness.

Imposing these two filters reduces the number of equations from which we derive the IA variable from 8,801 to 8,695 equations. Figure 7 demonstrates that the intraday variation of information asymmetry remains qualitatively unchanged from the intraday pattern of information asymmetry based on the whole 8,801 equations.

======Insert Figure 7 Here======

8.4 Robustness Test 4: Instrumental Variable

We test the relationship between the spread and the degree of information asymmetry and the relationship between the quote depth and the degree of information asymmetry by using IA as an independent variable. The variable IA is calculated from the regression results of Equation (1). Therefore, it might be related to the error term in Equation (7) and might lead to a bias in the regression results. We employ the rank of variable IA as an instrumental variable and run the regression again to see whether the relationships remain consistent. The test equations including the instrumental variable are as follows:

$$SPRD_{k,i} = \omega_0 + \omega_1 IA^*{}_{k,i} + \omega_2 VOL_{k,i} + \omega_3 VAR_{k,i} + \omega_4 PRC_{k,i} + \varepsilon_{k,i} \qquad (14.1)$$

$$Depth_{k,i} = \omega_0 + \omega_1 IA^*{}_{k,i} + \omega_2 VOL_{k,i} + \omega_3 VAR_{k,i} + \omega_4 PRC_{k,i} + \varepsilon_{k,i}, \qquad (14.2)$$

where the rank of variable *IA* is serving as an instrumental variable and $IA^*{}_{k,i}$ is the fitted variable obtained by running regression analysis of the *IA* on the instrumental variable: rank of IA and the estimation equation is: $IA_{k,i} = \alpha_0 + \alpha_1 rankofIA_{k,i} + \varepsilon_{k,i}$.

Table 19 reports the regression results of Equations (14.1) and (14.2). From the estimation of parameter ω_1, we find that the spread is positively related to the degree of information asymmetry and the quote depth is negatively related to the degree of information asymmetry. Therefore, we believe that the use of IA as an independent variable in Equation (7) does not result in a biased estimation of the relationships between the spread and the information asymmetry and between the depth and the information asymmetry.

======Insert Table 19 Here======

8.5 Robustness Test 5: Return Variance *VS* Return Volatility

In Equation (7), we used the return variance that measures the risk of stocks as one of the control variables. Using the variance as opposed to the standard deviation to proxy for risk might magnify the role of risk in determining the spread or the depth. As a robustness test, we repeat the same analysis using in the regression model the return volatility instead of the return variance. The test model is as follows:

$$SPRD_{k,i} = \tau_0 + \tau_1 IA_{k,i} + \tau_2 VOL_{k,i} + \tau_3 Volatility_{k,i} + \tau_4 PRC_{k,i} + \varepsilon_{k,i} \tag{15.1}$$

$$DPTH_{k,i} = \tau_0 + \tau_1 IA_{k,i} + \tau_2 VOL_{k,i} + \tau_3 Volatility_{k,i} + \tau_4 PRC_{k,i} + \varepsilon_{k,i}, \tag{15.2}$$

where $Volatility_{k,i}$ is the return volatility for stock k within trading interval i.

The regression result for Equations (15.1) and (15.2) are reported in Table 20. The relation between the spread and volatility is positive while the depth is negatively related to volatility. These relations are statistically significant at the 5% level or better. Similarly, the relationship between the spread and the degree of information asymmetry remains to be positive and significant and the relationship between the quote depth and the degree of information asymmetry remains to be negative and significant. Thus, the choice between the variance and the volatility does not change the qualitative conclusions of this thesis. We report the results based on variance to be consistent with previous studies.

======Insert Table 20 Here======

Chapter 9: Summary, Conclusions, and Directions for Future Research

This study measures the degree of information asymmetry over thirteen half-hour trading intervals within a trading day. We find that information asymmetry is highest at the beginning of the trading day and is lowest at the end of the trading day. Roughly, information asymmetry has a decreasing trend throughout the trading day. In particular, information asymmetry during intervals 1 to 6 is significantly higher than information asymmetry during trading interval 7. In contrast, information asymmetry during interval 13 is significantly lower than information asymmetry during interval 7.

Yet, an increase of information asymmetry from interval 7 to interval 10 is found to be significant. This behaviour might be interpreted as follows: First, the decline of the degree of information asymmetry is consistent with the argument that as trading goes on, information is gradually impounded into the transaction price and therefore the degree of information asymmetry should go down throughout the trading day. Second, the increase of the degree of information asymmetry right after the noon hour might be attributed to several factors. First, during the lunch break the rate of information arrival may stay constant but less information is resolved because of the lower trading activity. Therefore, there may be an accumulation of information that will be released right after the noon period. Second, there is a possibility that the rate of information arrival may increase after the noon hour. This increase may be the result of decisions taken by corporate boards or investment management committees that met during the morning and completed their meetings before and during the lunch hour. Third, the increase in information asymmetry might be attributed to the competition between better informed and less informed traders. Foster and Viswanathan (1994) suggest that the better informed traders strategically trade during earlier periods based on the common information available to both them and the less informed traders. When the common information has been resolved through periods of trading, the better informed traders will start to trade based on additional information only known by them.

Prior studies suggest that the spread and the depth are the two main tools that a market maker may employ to manage the information risk. Thus, we test the intraday variations of the spread and the depth. We find that the spread is widest at the open, and declines substantially during the early trading intervals and then levels off during the remaining trading hours. This behaviour is slightly different from the findings of previous studies that report that the spread changes following a U-shaped pattern. Instead of a U-shaped pattern our findings are consistent with an L-shaped pattern. The difference in the pattern at the end of a trading day might be attributed to the changes in the market that took place in the last decade.

In addition, the decoupling of the spread and the degree of information asymmetry suggests that the monopoly power of the market maker, proposed by Block and Kleidon (1992), may have more significant roles toward the end of the trading day. During the last trading interval the arrival of limit orders slows down (Chung ea al, 1999), which leaves the market maker greater monopoly power to keep the spread constant despite the reduction in adverse selection costs. Unfortunately, addressing these interesting issues goes beyond the scope of this paper and examining these propositions should be left to future studies.

For the quote depth, we observe that it is lowest at the open and then it increases consistently throughout the trading intervals. A comparison of the spread and the depth patterns suggests that the two change over the trading intervals in opposite directions. The results related to the degree of information asymmetry suggest that the depth and the spread change differently in response to changes in the degree of information asymmetry.

We examine empirically the relationship between the spread and the degree of information asymmetry and the relationship between the depth and the degree of information asymmetry across all the stocks in our data sample and across all trading intervals of a day. We find that the spread is positively related to the degree of information asymmetry while the depth is negatively related to it. This finding suggests that market makers will quote a wider spread or a lower depth, or both when greater informed trading is detected. In addition, we observe that the variations of the spread and the depth as we move forward from the morning trading interval are respectively positively and negatively related to the intraday variations in the degree of information asymmetry.

Collectively, we find that market makers face the greatest risk of information asymmetry during the opening interval and the least risk of information asymmetry during the last trading interval. Moreover when the information asymmetry is at the highest level during the open trading interval, the liquidity in the market is at the lowest both in the price dimension (the spread) and in the quantity dimension (the depth). As trading goes on, information is gradually resolved and correspondingly, the depth increases significantly toward the end of the trading day. That is to say, the quantity dimension of liquidity reaches the highest level during the last trading interval. However, the spread drops significantly from morning to noon and then levels off during the rest of the day, which implies that, on average, investors face higher trading costs during earlier trading period and then no significant difference in trading costs during the remaining trading intervals.

This study provides useful information to different participants in the market such as investors, financial professionals, policy makers and academics. First, our analysis of the intraday variations of the information asymmetry and the liquidity should help investors and financial professionals design better trading strategies. For example, liquidity traders might choose to trade later during a trading day since market liquidity increases as trading goes on. Second, through our study, policy makers can better understand market makers' quoting behaviour in order to build a fairly liquid stock market. Third, academically, we are the first to show how information asymmetry varies over the trading intervals by directly measuring the degree of information asymmetry and showing the relationship between the market liquidity and the information asymmetry in a dynamic context.

This study has limitations. We do not empirically test what may cause the degree of information asymmetry to rise following the noon hour. We can only speculate on the reasons. Analysis of this issue is beyond the scope of this study and may be pursued by future studies.

References

Affleck-Graves, J., Hedge, S. P., & Miller, R. E. (1994). Trading mechanisms and the components of the bid-ask spread. *The Journal of Finance*, 49(4), 1471-1488.

Amihud, Y., & Mendelson, H. (1982). Asset price behavior in a dealership market. *Financial Analysts Journal*, 38(3), 50-59.

Angel, J. J. (1997), Tick size, share prices, and stock splits. *The Journal of Finance*, 59(2), 655-681.

Ascioglu, A., Aydogdu, M., Chou, R. K., & Kugele, L. P. (2006). An analysis of intraday pattern in ETF and Common stock spreads. *Working paper*, Bryant University, National Central University and Christian Brothers University.

Bagehot, W. (1971). The only game in town. *The Financial Analysts Journal*, 27(2), 12-14, 22.

Benston, G. J., & Hagerman, R. L. (1974). Determinants of Bid-ask Spreads in the Over-the-counter Market. *Journal of Financial Economics*, 1(4), 353-364.

Brock, W. A., & Kleidon, A. W. (1992). Periodic market closure and trading volume: A model of intraday bids and asks. *Journal of Economic Dynamics and Control*, 16(3-4), 451-489.

Conroy, R. M., Harris, R. S., & Benet, B. A. (1990). The effects of stock splits on bid-ask spreads. *The Journal of Finance*, 45(4), 1285-1295.

Chakravarty, S., Van Ness, B. F., & Van Ness, R. A. (2005). The effect of decimalization on trade size and adverse selection costs. *Journal of Business Finance Accounting*, 32(5-6), 1063-1081.

Chakravarty, S., Wood, R. A., & Van Ness, R. A. (2004). Decimals and liquidity: A study of the NYSE. *The Journal of Financial Research*. 27(1), 75-94.

Chan, K. C., Christie, W.G., & Schultz, P. H. (1995). Market structure and the intraday pattern of bid-ask spreads for NASDAQ securities. *The Journal of Business*, 68(1), 35-60.

Choi, J. Y., Salandro, D., & Shastri, K. (1988). On the estimation of bid-ask spreads: Theory and evidence. The Journal of Financial and Quantitative Analysis, 23(2), 219-230.

Chung, K. H., & Van Ness, R. A. (2001). Order handling rules, tick size, and the intraday pattern of bid–ask spreads for Nasdaq stocks. *Journal of Financial Markets*, 4(2), 143-161.

Chung, K. H., Van Ness, B. F., & Van Ness, R. A. (1999). Limit orders and the bid–ask spread. *Journal of Financial Economics*, 53(2), 255-287.

Chung, K. H., & Zhao, X. (2003). Intraday variation in the bid-ask spread: Evidence after the market reform. *The Journal of Financial Research*, 26(2), 191-206.

Dupont, D. (2000). Market making, prices, and quantity limits. *The Review of Financial Studies*, 13(4), 1129-1151.

Easley, D. & O'Hara M. (1987) Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1), 69-90.

Foster, F. D., & Viswanathan, S. (1990). A Theory of the interday variation in volume, variance, and trading costs in securities markets. *The Review of Financial Studies*, 3(4), 593-624.

Foster, F. D., & Viswanathan, S. (1993). Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *The Journal of Finance*, 48(1), 187-211.

Foster, F. D., & Viswanathan, S. (1994). Strategic trading with asymmetrically informed traders and long-lived information. *The Journal of Financial and Quantitative Analysis*, 29(4), 499-518.

George, T. J., Kaul, G., & Nimalendran, M. (1991). Estimation of the bid-ask spread and its components: A new approach. *The Review of Financial Studies*, 4(4), 623-656.

Glosten, R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71-100.

Glosten, L. R. (1989). Insider trading, liquidity and the role of the monopolist specialist. *Journal of Business*, 62(2), 211-235.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21(1), 123-142.

Gong, N. (2007). Effectiveness and market reaction to the stock exchange's inquiry in Australia. *Journal of Business Finance Accounting*, 34(7-8), 1141-1168.

Greene, W. H. (2003). *Econometric analysis*, *5th edition,* Prentice Hall.

Gregoriou, A., Ioannidis, C., & Skerratt, L. (2005). Information asymmetry and the bid-ask spread: Evidence from the UK. *Journal of Business Finance Accounting*, 32(9-10), 1801-1826.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators.

Econometrica, 50(4), 1029-1054.

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics*, 22(2), 229-252.

Hasbrouck, J, Sofianos, G., & Sosebee, D. (1993). New York Stock Exchange systems and trading procedures. Working paper, *New York University* and *NYSE*.

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. *The Review of Financial Studies*, 6(3), 473-506.

Heflin, F., Shaw, K. W., & Wild, J. J. (2007). Disclosure policy and intraday spread patterns. *The Review of Accounting and Finance*, 6(3), 285-303.

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies*, 10(4), 995-1034.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315-1335.

Kennedy, P. (2003). *A guide to econometrics*, *5th edition*, the MIT Press Cambridge, Massachusetts..

Lee, C. M. C., Mucklow, B., & Ready, M. J. (1993). Spreads, depths, and the impact of earnings information: An intraday analysis. *The Review of Financial Studies*, 6(2), 345-374.

Lee, C. M. C., & Ready, M. J. (1991). Inferring trade direction from intraday data. *The Journal of Finance*, 46(2), 733-746.

Li, Y., Van Ness, B. F., & Van Ness, R. A. (2005). Daily and intraday patterns in spread and depth for limit orders and specialists. *Quarterly Journal of Business and Economics*, 44(3/4), 3-14.

Lin, J. C., Sanger, G. C., & Booth, G. G. (1995). Trade size and components of the bid-ask spread. *The Review of Financial Studies*, 8(4), 1153-1183.

Lipson, M. (2001). Stock splits, liquidity and limit orders. Working paper, *University of Georgia*.

Madhavan, A. (1992). Trading mechanisms in securities markets. *The Journal of Finance*, 47(2), 607-641.

Madhavan, A., Richardson, M., & Roomans, M. (1997). why do security prices fluctuate? A transaction-level analysis of NYSE stocks. *The Review of Financial Studies*, 10, 1035-1064.

Madhavan, A. and S. Smidt (1991). A Bayesian Model of Intraday Specialist Pricing. *The*

Journal of Financial Economics, 30(1), 99-134.

McInish, T. H., & Van Ness, B. F. (2002). An intraday examination of the components of the bid-ask spread. *The Financial Review*, 37(4), 507-524.

McInish, T. H., & Wood, R. A. (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *The Journal of Finance*, 47(2), 753-764.

Noronha, G.M., Sarin, A., & Saudagaran, S. M. (1996). Testing for microstructure effects of international dual listings using intraday data. *Journal of Banking Finance*, 20(6), 965-983.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4), 1127-1139.

SAS/ETS User's Guide, Version9, SAS.

Schultz, P. (2000). Stock splits, tick size, and sponsorship. *The Journal of Finance*, 55(1), 429-450.

Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance*, 33(4), 1133-1151.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *The Journal of Finance*, 44(1), 115-134.

Tannous, F. T., & Zhang, Y. (2008). Cross-listing and trading on the domestic market: Evidence from Canada-US partial holidays. *Journal of Business Finance Accounting*, 35(9-10), 1245.

Tinic, S. M. (1972). The economics of liquidity services. *The Quarterly Journal of Economics*, 86(1), 79-93.

Van Ness, B. F., Van Ness, R. A., & Warr, R. S. (2001). Determinants of adverse selection on the NYSE and Nasdaq: A paired comparison. Working paper, *University of Mississippi* and *North Carolina State University*.

Van Ness, B. F., Van Ness, R. A., & Warr, R. S. (2002). Is the adverse selection component real higher on the NYSE/Amex than on the Nasdaq? *Journal of Business Finance Accounting*, 29(5-6), 807-824.

Zhao, X., & Chung, K. H. (2006). Decimal pricing and information-based trading: Tick size and informational efficiency of asset price. *Journal of Business Finance Accounting*, 33(5-6), 753-766.

Tables and Figures

Table 1: Summary of the Literature

Components of the Spread
OP cost: Benston and Hagerman (1974) and Stoll (1978) IH cost: Tinic (1972) and Stoll (1978) AI cost: Bagehot(1971) Empirically, Stoll (1989), Glosten and Harris's (1988) and Hasbrouk (1988) showed the AI cost constitutes a significant component of the spread.

Intraday Pattern of the Spread
U-shaped Pattern (Evidence from NYSE): Brock and Kleidon (1992), McInish and Wood (1992), Lee et al (1993), Madhavan (1997), McInish and Van Ness (2002) and Heflin (2007) Ascioglu et al (2006) showed that the spread drops sharply during early trading hours and then remain flat during the rest of trading day which is different from previous studies. Explanations of the Intraday Pattern of the Spread Market maker power: Block and Kleidon (1992) Inventory control effect: Amihud and Mendelson (1982) Information asymmetry effect: Kyle (1985), Madhavan (1992) and Foster and Viswanathan (1994)

Intraday Pattern of the Asymmetric Information Component
Decreasing over the trading day (AI cost): Foster and Viswanathan (1993), Madhavan et al (1997) and McInish and Van Ness (2002)

Relationship between the Spread and the Information Asymmetry and between the Depth and the Information Asymmetry
McInish and Wood (1992), Gregoriou et al (2005), Gong (2007) found that Market makers widen the spread due to higher level of information asymmetry. Lee et al (1993) and Dupont (2000) indicated that the spread will be wider and the quote depth will be lower while the information asymmetry is higher.

Table 2: Summary Statistics for the Sample of NYSE Stocks

Variable	Mean	Min.	Max.	Std. Dev.
Share price($ per share)	36.16	4.98	153.59	20.92
Share volume (million)	240.79	0.113	4,497.67	410.41
Trade size (shares)	539.10	144.64	3671.61	399.23
Number of trades (per interval)	25,736.78	16.00	157,528.00	21,798.32

Note: Share price is the transaction price per trade, share volume is the total number of shares traded for a year, trade size is the number of shares traded per transaction and number of trades is the number of transactions in each trading interval. The sample consists of 677 stocks and the share price, the share volume and the trade size are first computed by each stock through the whole year, and then the Table entries are calculated from all the stock-specific data. The number of trades is first counted for each trading interval of each stock and then the Table entries are calculated across all the 13 trading intervals of all 677 stocks.

Table 3: The Frequency of Positive, Negative, Significant, or Insignificant Coefficients of the Independent Variables in Equation (1)

Panel A: Results from the OLS Estimation

	Coefficient: β_1		Coefficient: β_2		Coefficient: β_3	
Significant at 0.01	Negative	Positive	Negative	Positive	Negative	Positive
Significant (%)	22 (0.25%)	8145 (92.55%)	0 (0%)	8741 (99.32%)	8676 (98.58%)	0 (0%)
Not significant (%)	72 (0.82%)	562 (6.39%)	10 (0.11%)	50 (0.57%)	108 (1.23%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.05						
Significant (%)	25 (0.28%)	8342 (94.78%)	0 (0 %)	8761 (99.55%)	8714 (99.01%)	0 (0%)
Not significant (%)	69 (0.78%)	365 (4.15%)	10 (0.11%)	30 (0.34%)	70 (0.80%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.1						
Significant (%)	29 (0.33%)	8416 (95.63%)	1 (0.01%)	8766 (99.60%)	8734 (99.24%)	0 (0%)
Not significant (%)	65 (0.74%)	291 (3.31%)	9 (0.10%)	25 (0.28%)	50 (0.57%)	17 (0.19%)
Total	94	8707	10	8791	8784	17

$\beta_2 > |\beta_3|$ Restriction : 8771 equations satisfy this size restriction

Note: Panel A shows the regression results for Equation (1) based on the OLS estimation. The number of positive and negative estimates of independent variables of Equation (1) and the number and percentage of significant and insignificant estimates are reported respectively at the 0.01, 0.05, and 0.1 levels based on a two-tailed test. All the percentages are computed based on the total number of equation results equal to 8,801. At the bottom, we also report the number of equations, among the 8,801, that satisfy the restriction: $\beta_2 > |\beta_3|$.

Table 3: The Frequency of Positive, Negative, Significant, or Insignificant Coefficients of the Independent Variables in Equation (1)

Panel B: Results from the GMM Estimation

	Coefficient: β_1		Coefficient: β_2		Coefficient: β_3	
Significant at 0.01	Negative	Positive	Negative	Positive	Negative	Positive
Significant (%)	0 (0%)	5904 (67.08%)	0 (0%)	8736 (99.26%)	8671 (98.52%)	0 (0%)
Not significant (%)	94 (1.07%)	2803 (31.85%)	10 (0.11%)	55 (0.62%)	113 (1.28%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.05						
Significant (%)	2 (0.02%)	7130 (81.01%)	1 (0.01%)	8762 (99.56%)	8709 (98.95%)	0 (0%)
Not significant (%)	92 (1.05%)	1577 (17.92%)	9 (0.10%)	29 (0.33%)	75 (0.85%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.1						
Significant (%)	4 (0.05%)	7649 (86.91%)	1 (0.01%)	8765 (99.59%)	8727 (99.16%)	0 (0%)
Not significant (%)	90 (1.02%)	1058 (12.02%)	9 (0.10%)	26 (0.30%)	57 (0.65%)	17 (0.19%)
Total	94	8707	10	8791	8784	17

$\beta_2 > |\beta_3|$ Restriction : 8771 equations satisfy this size restriction

Note: Panel B shows the regression results for Equation (1) based on GMM estimation. The number of positive and negative estimates of independent variables of Equation (1) and the number and percentage of significant and insignificant estimates are reported respectively at the 0.01, 0.05, and 0.1 levels based on a two-tailed test. All the percentages are computed based on the total number of equations being 8,801. At the bottom, we also report the number of equations, among the 8,801, that satisfy the restriction: $\beta_2 > |\beta_3|$.

Table 4: Summary Statistics of IA Calculated from Equation (1) and Intraday Pattern Test for IA

Panel A: Descriptive Statistics for IA = 1 – PRIOR

Interval	N	Lower Bound	Mean	Upper Bound
9:30am-10:00am Interval 1	677	0.5741	0.5873	0.6004
10:00 am-10:30 am Interval 2	677	0.5258	0.5340	0.5423
10:30 am-11:00 am Interval 3	677	0.5124	0.5250	0.5375
11:00 am-11:30 am Interval 4	677	0.4956	0.5054	0.5153
11:30 am-12:00 pm Interval 5	677	0.4851	0.4998	0.5146
12:00 pm-12:30 pm Interval 6	677	0.4846	0.4949	0.5051
12:30 pm-1:00 pm Interval 7	677	0.4660	0.4765	0.4870
1:00 pm-1:30 pm Interval 8	677	0.4725	0.4829	0.4934
1:30 pm-2:00 pm Interval 9	677	0.4754	0.4892	0.5030
2:00 pm-2:30 pm Interval 10	677	0.4923	0.5070	0.5216
2:30 pm-3:00 pm Interval 11	677	0.4608	0.4767	0.4927
3:00 pm-3:30 pm Interval 12	677	0.4636	0.4755	0.4873
3:30 pm-4:00 pm Interval 13	677	0.4001	0.4107	0.4213

Note: In Panel A, all the descriptive statistics are calculated based on variable IA calculated from the PRIOR values, which are obtained from the regression results of Equation (1). The lower bound and upper bound are for a 95% confidence interval.

Table 4: Summary Statistics of IA Calculated from Equation (1) and Intraday Pattern Test for IA

Panel B: Estimated Coefficients of Equation (2)				
Interval	Variable	Estimates	t-statistics	p-value
	Intercept	0.4765***	76.34	p < 0.0001
9:30am-10:00am Interval 1	DUM1	0.1108***	12.55	p < 0.0001
10:00 am-10:30 am Interval 2	DUM2	0.0575***	6.52	p < 0.0001
10:30 am-11:00 am Interval 3	DUM3	0.0485***	5.49	p < 0.0001
11:00 am-11:30 am Interval 4	DUM4	0.0289***	3.28	p = 0.001
11:30 am-12:00 pm Interval 5	DUM5	0.0233***	2.64	p = 0.0082
12:00 pm-12:30 pm Interval 6	DUM6	0.01848**	2.08	p = 0.0374
1:00 pm-1:30 pm Interval 8	DUM8	0.0064	0.73	p = 0.4653
1:30 pm-2:00 pm Interval 9	DUM9	0.0127	1.44	p = 0.1501
2:00 pm-2:30 pm Interval 10	DUM10	0.0305***	3.45	p = 0.0006
2:30 pm-3:00 pm Interval 11	DUM11	0.0002	0.03	p = 0.9789
3:00 pm-3:30 pm Interval 12	DUM12	-0.0010	-0.12	p = 0.9084
3:30 pm-4:00 pm Interval 13	DUM13	-0.0658***	-7.45	p < 0.0001

*** and ** respectively indicate significance at the 0.01 level and the 0.05 level.

Table 5: Summary Statistics of SPRD Calculated from Equation (3) and Intraday Pattern Test for SPRD

Panel A: Summary Statistics of SPRD (time-weighted average relative spread)				
Interval	N	Lower Bound	Mean	Upper Bound
9:30am-10:00am				
Interval 1	677	0.002025	0.002199	0.002372
10:00 am-10:30 am				
Interval 2	677	0.001431	0.001569	0.001706
10:30 am-11:00 am				
Interval 3	677	0.001252	0.001380	0.001508
11:00 am-11:30 am				
Interval 4	677	0.001160	0.001276	0.001391
11:30 am-12:00 pm				
Interval 5	677	0.001124	0.001243	0.001361
12:00 pm-12:30 pm				
Interval 6	677	0.001113	0.001237	0.001361
12:30 pm-1:00 pm				
Interval 7	677	0.001085	0.001209	0.001332
1:00 pm-1:30 pm				
Interval 8	677	0.001045	0.001162	0.001279
1:30 pm-2:00 pm				
Interval 9	677	0.001063	0.001191	0.001320
2:00 pm-2:30 pm				
Interval 10	677	0.001050	0.001156	0.001262
2:30 pm-3:00 pm				
Interval 11	677	0.001032	0.001145	0.001259
3:00 pm-3:30 pm				
Interval 12	677	0.000994	0.001099	0.001205
3:30 pm-4:00 pm				
Interval 13	677	0.000998	0.001111	0.001224

Note: In panel A, SRPD in Equation (3) is first calculated for each stock during each trading interval and then the descriptive statistics of SRPD for each interval is computed based on the 677 calculations of SRPD for each interval. The lower bound and upper bound are for a 95% confidence interval.

Table 5: Summary Statistics of SPRD Calculated from Equation (3) and Intraday Pattern Test for SPRD

Panel B: Estimated Coefficients of Equation (4)				
Interval	Variable	Estimates	t-statistics	p-value
	Intercept	0.001209***	19.07	$p < 0.0001$
9:30am-10:00am Interval 1	DUM1	0.000990***	11.04	$p < 0.0001$
10:00 am-10:30 am Interval 2	DUM2	0.000360***	4.01	$p < 0.0001$
10:30 am-11:00 am Interval 3	DUM3	0.000171	1.91	$p = 0.0561$
11:00 am-11:30 am Interval 4	DUM4	0.000067	0.74	$p = 0.4565$
11:30 am-12:00 pm Interval 5	DUM5	0.000034	0.38	$p = 0.7073$
12:00 pm-12:30 pm Interval 6	DUM6	0.000028	0.32	$p = 0.7515$
1:00 pm-1:30 pm Interval 8	DUM8	-0.000047	-0.52	$p = 0.6024$
1:30 pm-2:00 pm Interval 9	DUM9	-0.000017	-0.19	$p = 0.8460$
2:00 pm-2:30 pm Interval 10	DUM10	-0.000053	-0.59	$p = 0.5542$
2:30 pm-3:00 pm Interval 11	DUM11	-0.000063	-0.71	$p = 0.4795$
3:00 pm-3:30 pm Interval 12	DUM12	-0.000109	-1.22	$p = 0.2224$
3:30 pm-4:00 pm Interval 13	DUM13	-0.000098	-1.09	$p = 0.2759$

*** indicates significance at the 0.01 level.

Table 6: Summary Statistics of DPTH Calculated from Equation (5) and Intraday Pattern Test for DPTH

Panel A: Summary statistics of DPTH (unit: 100 shares)				
Interval	N	Lower Bound	Mean	Upper Bound
9:30am-10:00am				
Interval 1	677	7.21	7.88	8.55
10:00 am-10:30 am				
Interval 2	677	7.80	8.62	9.45
10:30 am-11:00 am				
Interval 3	677	8.13	9.07	10.00
11:00 am-11:30 am				
Interval 4	677	8.38	9.38	10.38
11:30 am-12:00 pm				
Interval 5	677	8.49	9.50	10.51
12:00 pm-12:30 pm				
Interval 6	677	8.54	9.57	10.61
12:30 pm-1:00 pm				
Interval 7	677	8.61	9.68	10.74
1:00 pm-1:30 pm				
Interval 8	677	8.89	10.00	11.12
1:30 pm-2:00 pm				
Interval 9	677	9.04	10.17	11.29
2:00 pm-2:30 pm				
Interval 10	677	9.18	10.29	11.41
2:30 pm-3:00 pm				
Interval 11	677	9.39	10.53	11.67
3:00 pm-3:30 pm				
Interval 12	677	10.28	11.49	12.70
9:30am-10:00am				
Interval 13	677	13.35	14.77	16.19

Note: In panel A, DPTH in Equation (5) is first calculated for each stock during each trading interval, and then the descriptive statistics are computed based on the 677 calculations of DPTH for each interval. The lower bound and upper bound are for a 95% confidence interval.

Table 6: Summary Statistics of DPTH Calculated from Equation (5) and Intraday Test for DPTH

Panel B: Estimated Coefficients of Equation (6)				
Interval	Variable	Estimates	t-statistics	p-value
	Intercept	9.6751***	17.84	p < 0.0001
9:30am-10:00am Interval 1	DUM1	-1.7937**	-2.34	p = 0.0194
10:00 am-10:30 am Interval 2	DUM2	-1.0502	-1.37	p = 0.1711
10:30 am-11:00 am Interval 3	DUM3	-0.6079	-0.79	p = 0.4282
11:00 am-11:30 am Interval 4	DUM4	-0.2968	-0.39	p = 0.6988
11:30 am-12:00 pm Interval 5	DUM5	-0.1703	-0.22	p = 0.8243
12:00 pm-12:30 pm Interval 6	DUM6	-0.1002	-0.13	p = 0.8961
1:00 pm-1:30 pm Interval 8	DUM8	0.3298	0.43	p = 0.6672
1:30 pm-2:00 pm Interval 9	DUM9	0.4931	0.64	p = 0.5204
2:00 pm-2:30 pm Interval 10	DUM10	0.6165	0.80	p = 0.4216
2:30 pm-3:00 pm Interval 11	DUM11	0.8514	1.11	p = 0.2671
3:00 pm-3:30 pm Interval 12	DUM12	1.8187**	2.37	p = 0.0178
3:30 pm-4:00 pm Interval 13	DUM13	5.0979***	6.65	p < 0.0001

*** and ** respectively indicate significance at the 0.01 level and the 0.05 level.

Table 7: Descriptive Statistics for Independent Variables of Equation (7)

Mean Values by trading interval					
Interval	$SPRD_{k,i}$	$VOL_{k,i}$	$VAR_{k,i}$	$PRICE_{k,i}$	$IA_{k,i}$
9:30am-10:00am					
Interval 1	0.002199	23,235,994	0.000001917	36.21	0.5873
10:00 am-10:30am					
Interval 2	0.001569	22,740,971	0.000001321	36.16	0.5340
10:30 am-11:00 am					
Interval 3	0.001380	20,040,198	0.000000999	36.15	0.5250
11:00 am-11:30 am					
Interval 4	0.001276	17,478,278	0.000001057	36.18	0.5054
11:30 am-12:00 pm					
Interval 5	0.001243	15,283,861	0.000000872	36.16	0.4998
12:00 pm-12:30 pm					
Interval 6	0.001237	13,908,849	0.000000853	36.16	0.4949
12:30 pm-1:00 pm					
Interval 7	0.001209	12,549,695	0.000000887	36.17	0.4765
13:00 pm-13:30 pm					
Interval 8	0.001162	12,892,328	0.000000791	36.13	0.4829
1:30 pm-2:00 pm					
Interval 9	0.001191	13,808,037	0.000000816	36.14	0.4892
2:00 pm-2:30 pm					
Interval 10	0.001156	15,947,021	0.000000799	36.13	0.5070
2:30 pm-3:00 pm					
Interval 11	0.001145	17,407,701	0.000000802	36.14	0.4767
3:00 pm-3:30 pm					
Interval 12	0.001099	21,231,634	0.000000791	36.13	0.4755
3:30 pm-4:00 pm					
Interval 13	0.001111	34,266,939	0.000000774	36.16	0.4107

Note: This Table reports the mean values of SPRD, trading volume, return variance, average share price and IA for each of the 13 trading intervals. First, all the variables are calculated for each stock within each trading interval as defined in Equation (7). Then the mean values of them are computed by interval across all the 677 stocks.

Table 8: Estimated Coefficients of the Regression Equation (9)

Dependent Variable: $\widehat{SPRD}$				
Independent Variable	$\widehat{IA}$	$\widehat{VOL}$	$\widehat{VAR}$	$\widehat{PRC}$
Estimated Coefficients	0.000813***	1.32E-12***	161.2889***	0.000129***
	(22.30)	(3.40)	(53.73)	(4.65)
	($p < 0.0001$)	($p = 0.0007$)	($p < 0.0001$)	($p < 0.0001$)
Adjusted R^2	0.9333			
F test for no fixed effects	(36.90)			
	($p < 0.0001$)			

*** indicate significance at the 0.01 level.

Note: Table 8 shows the parameter estimates for Equation (9), including the independent variables: information asymmetry, trading volume, return variance and price. All those variables are in deviation forms as defined in Equation (9). The numbers in parentheses are the t-statistics and p-values. The adjusted R^2 is also reported. In the column at the end, the F-test for no fixed effect is shown and in parentheses are the F-statistic and the corresponding p-value.

Table 9: The Correlations among the Independent Variables of Equation (9)

	$\widehat{IA}$	$\widehat{VOL}$	$\widehat{VAR}$	$\widehat{PRC}$
$\widehat{IA}$	1	-0.0443*** (p < 0.0001)	0.0216*** (p = 0.0424)	-0.0353 (p = 0.0009)
$\widehat{VOL}$	-0.0443*** (p < 0.0001)	1	0.0033 (p = 0.7553)	0.0058 (p = 0.5861)
$\widehat{VAR}$	0.0216** (p = 0.0424)	0.0033 (p = 0.7553)	1	0.0025 (p = 0.8160)
$\widehat{PRC}$	-0.0353*** (p = 0.0009)	0.0058 (p = 0.5861)	0.0025 (p = 0.8160)	1

*** and ** respectively indicate significance at the 0.01 level and the 0.05 level.

Note: Table 9 offers the correlation matrix, which examines the correlation among the independent variables of Equation (9) and all the variables here are the value of deviations from group mean defined in Equation (9). The Pearson correlation coefficient for Each pair of the 4 variables is reported here. The numbers in parentheses are p-values.

Table 10: Cross Sectional Descriptive Statistics for Variables in Equation (10)

Variable	N	Mean	Min	Max	Std
SPRD	677	0.001306	0.0002578	0.01484	0.001599
IA	677	0.4973	0.1161	0.7857	0.1089
VOL	677	18,522,423.52	8,707.69	345,974,892.00	31,570,476.36
VAR	677	0.0000009754	0.00000001952	0.00003272	0.000003003
PRC	677	36.16	4.99	153.76	20.92

Note: All the variables in Table 10 are as defined in Equation (10). For each stock during each trading interval, there is one calculated value. To show the cross sectional variation, all the variables for each stock are first averaged across the 13 trading intervals and the descriptive statistics are then computed using the stock-specific data.

Table 11: Estimated Coefficients of the Regression Equation (10)

Dependent Variable: SPRD			
Independent Variable	Estimated Coefficients	t-Statistics	p-Value
Intercept	0.001786***	39.75	$p < 0.0001$
IA	0.000396***	7.10	$p < 0.0001$
VOL	-4.65E-12***	-17.69	$p < 0.0001$
VAR	399.53***	159.62	$p < 0.0001$
PRC	-0.000013***	-31.18	$p < 0.0001$
DUM2	-0.000374***	-9.06	$p < 0.0001$
DUM3	-0.000443***	-10.70	$p < 0.0001$
DUM4	-0.000574***	-13.82	$p < 0.0001$
DUM5	-0.000542***	-13.00	$p < 0.0001$
DUM6	-0.000544***	-13.03	$p < 0.0001$
DUM7	-0.000585***	-13.94	$p < 0.0001$
DUM8	-0.000595***	-14.19	$p < 0.0001$
DUM9	-0.000573***	-13.71	$p < 0.0001$
DUM10	-0.000599***	-14.40	$p < 0.0001$
DUM11	-0.000592***	-14.16	$p < 0.0001$
DUM12	-0.000616***	-14.75	$p < 0.0001$
DUM13	-0.000511***	-12.09	$p < 0.0001$
Adjusted R^2: 0.7949			

*** indicates significance at the 0.01 level.

Note: All the variables in Table 11 are defined in Equation (10). The estimated coefficients for each variable in the Equation (10) is displayed including the corresponding the t-statistics and the p-values. At the bottom, the adjusted R^2 is reported for the regression of Equation (10).

Table 12: Descriptive Statistics for Independent Variables of Equation (11)

Mean Values by trading interval					
Interval	$DPTH_{k,i}$	$VOL_{k,i}$	$VAR_{k,i}$	$PRICE_{k,i}$	$IA_{k,i}$
9:30am-10:00am					
Interval 1	7.88	23,235,994	0.000001917	36.21	0.5873
10:00 am-10:30am					
Interval 2	8.62	22,740,971	0.000001321	36.16	0.5340
10:30 am-11:00am					
Interval 3	9.07	20,040,198	0.000000999	36.15	0.5250
11:00 am-11:30am					
Interval 4	9.38	17,478,278	0.000001057	36.18	0.5054
11:30 am-12:00pm					
Interval 5	9.50	15,283,861	0.000000872	36.16	0.4998
12:00 pm-12:30pm					
Interval 6	9.57	13,908,849	0.000000853	36.16	0.4949
12:30 pm-1:00 pm					
Interval 7	9.68	12,549,695	0.000000887	36.17	0.4765
1:00 pm-1:30 pm					
Interval 8	10.00	12,892,328	0.000000791	36.13	0.4829
1:30 pm-2:00 pm					
Interval 9	10.17	13,808,037	0.000000816	36.14	0.4892
2:00 pm-2:30 pm					
Interval 10	10.29	15,947,021	0.000000799	36.13	0.5070
2:30 pm-3:00 pm					
Interval 11	10.53	17,407,701	0.000000802	36.14	0.4767
3:00 pm-3:30 pm					
Interval 12	11.49	21,231,634	0.000000791	36.13	0.4755
9:30am-10:00am					
Interval 13	14.77	34,266,939	0.000000774	36.16	0.4107

Note: Table 12 is the same as Table 7, except that the SPRD column is replaced by the mean values of DPTH for each interval of the 13 trading intervals, obtained from Table 6. The unit of DPTH is 100 shares.

Table 13: Estimated Coefficients of Equation (11)

Dependent Variable: $\widehat{DPTH}$				
Independent Variable	$\widehat{IA}$	$\widehat{VOL}$	$\widehat{VAR}$	$\widehat{PRC}$
Estimated Coefficients	-3.1639*** (-12.14) ($p < .0001$)	1.02E-7*** (36.50) ($p < .0001$)	-106927*** (-4.98) ($p < .0001$)	-0.6941*** (-3.49) ($p = .0005$)
Adjusted R^2	0.9686			
F test for no fixed effects	(150.73) ($p < .0001$)			

*** indicates significance at the 0.01 level.

Note: All variables including dependent and independent variables are the deviations from the group mean as defined in Equation (11). The estimated coefficients of all the independent variables are reported here. The numbers in parentheses are the t-statistics and p-values. The adjusted R^2 value and the F-statistic for the fixed effects test are also shown in Table 13.

Table 14: Estimated Coefficients of Equation (12)

Dependent Variable: DPTH

Independent Variable	Estimated Coefficients	t-Statistics	p-Value
Intercept	14.14***	24.41	$p < 0.0001$
IA	-11.46***	-15.92	$p < 0.0001$
VOL	2.69E-07***	79.36	$p < 0.0001$
VAR	-47699	-1.48	$p = 0.1394$
PRC	-0.1572***	-28.79	$p < 0.0001$
DUM2	0.2291	0.43	$p = 0.6671$
DUM3	1.2777**	2.39	$p = 0.0167$
DUM4	2.0612***	3.85	$p < 0.0001$
DUM5	2.7032***	5.03	$p < 0.0001$
DUM6	3.0852***	5.73	$p < 0.0001$
DUM7	3.3433***	6.18	$p < 0.0001$
DUM8	3.6440***	6.75	$p < 0.0001$
DUM9	3.6350***	6.74	$p < 0.0001$
DUM10	3.3846***	6.31	$p < 0.0001$
DUM11	2.8822***	5.35	$p < 0.0001$
DUM12	2.8034***	5.21	$p < 0.0001$
DUM13	1.8365***	3.37	$p = 0.0007$

Adjusted R^2: 0.5268

*** and ** indicates significance at the 0.01 and the 0.05 level respectively

Note: All the variables in Table 14 are all defined in Equation (12), the estimated coefficients of all the independent variables are reported in Table 14 and the corresponding t-statistics and p-values are also reported. At the bottom is the adjusted R^2 of the regression.

Table 15: Correlations between SPRD and DPTH for each of the 13 Intervals

Interval	Correlations	p-values
9:30am-10:00am		
Interval 1	-0.1310***	p = 0.0006
10:00 am-10:30 am		
Interval 2	-0.1015***	p = 0.0082
10:30 am-11:00 am		
Interval 3	-0.0872**	p = 0.0233
11:00 am-11:30 am		
Interval 4	-0.0698*	p = 0.0696
11:30 am-12:00 pm		
Interval 5	-0.0643*	p = 0.0944
12:00 pm-12:30 pm		
Interval 6	-0.0512	p = 0.1833
12:30 pm-1:00 pm		
Interval 7	-0.0528	p = 0.1703
1:00 pm-1:30 pm		
Interval 8	-0.0499	p = 0.1945
1:30 pm-2:00 pm		
Interval 9	-0.0410	p = 0.2872
2:00 pm-2:30 pm		
Interval 10	-0.0458	p = 0.234
2:30 pm-3:00 pm		
Interval 11	-0.0606	p = 0.1152
3:00 pm-3:30 pm		
Interval 12	-0.0608	p = 0.1142
9:30am-10:00am		
Interval 13	-0.0981**	p = 0.0107

***, ** and * indicates significance at 0.01, 0.05 and 0.1 level respectively

Note: The Pearson correlations between the spread and the depth for each of the 13 intervals are reported here. The calculations of the spread and the depth are respectively defined in Equations (3) and (5). For each trading interval, there are 677 observations for both spread and depth.

Table 16: The Frequency of Positive, Negative, Significant, or Insignificant Coefficients of the Independent Variables in Equation (1)

Results Based on different Bandwidths for each Regression (GMM)						
	Coefficient: β_1		Coefficient: β_2		Coefficient: β_3	
Significant at 0.01	Negative	Positive	Negative	Positive	Negative	Positive
Significant (%)	0 (0%)	5894 (66.97%)	0 (0%)	8734 (99.24%)	8665 (98.45%)	0 (0%)
Not significant (%)	94 (1.07%)	2813 (31.96%)	10 (0.11%)	57 (0.65%)	119 (1.35%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.05						
Significant (%)	1 (0.01%)	7121 (80.91%)	0 (0 %)	8755 (99.48%)	8703 (98.89%)	0 (0%)
Not significant (%)	93 (1.06%)	1586 (18.02%)	10 (0.11%)	36 (0.41%)	81 (0.92%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
Significant at 0.1						
Significant (%)	3 (0.03%)	7650 (86.92%)	0 (0%)	8764 (99.58%)	8725 (99.14%)	0 (0%)
Not significant (%)	91 (1.03%)	1057 (12.01%)	10 (0.11%)	27 (0.31%)	59 (0.67%)	17 (0.19%)
Total	94	8707	10	8791	8784	17
$\beta_2 > \lvert\beta_3\rvert$ Restriction : 8771 equations satisfy this size restriction						

Note: Table 16 reports the regression results for Equation (1) based on GMM estimation with different bandwidths for each regression. The number of positive and negative estimates of independent variables of Equation (1) and the number and percentage of significant and insignificant estimates are reported respectively at the 0.01, 0.05, and 0.1 levels based on the two-tailed test. All the percentages are computed based on the total number of equation results equal to 8,801. At the bottom, we also report the number of equations, among the 8,801, that satisfy the restriction: $\beta_2 > |\beta_3|$.

Table 17: Summary Statistics for Three Groups of Stocks

	High activity group (N = 225)	Medium activity group (N = 227)	Low activity group (N = 225)
IA	0.4376	0.5141	0.5401
SPRD	0.000576	0.000872	0.002474
DPTH	17.59	7.17	5.48

t-test for comparing the means

	High versus medium	High versus low	Medium versus low
IA	t = 2.06(p = .0001) ***	t = 2.07(p < .0001) ***	t = 2.06(p = .1033)
SPRD	t = 2.10(p < .002)***	t = 2.14(p < .0001) ***	t = 2.10(p < .0001) ***
DPTH	t = 2.11(p < .0001)***	t = 2.16(p < .0001) ***	t = 2.09(p = .0004)***

*** indicates significance at the 0.01 level

Note: The average values of variables: IA, SPRD, and DPTH are reported in Table 17. All the variables are already defined and calculated above. The average values are computed by first averaging the interval values of each variable across the stocks within each group and then calculating the mean value of each variable across the 13 trading intervals for each group. For each of the variables: IA, SPRD, and DPTH, the t-test is conducted respectively for the high activity group versus medium activity group, the high activity group versus low activity group and the medium activity group versus low activity group. The t-statistics are reported and the numbers in parentheses are the p-values.

Table 18: Estimated Results for Equation (13.1) to Equation (13.6)

	SPRD	DPTH
Independent Variables		
High Activity		
Intercept	0.000443***	42.01***
	(51.05)	(31.75)
IA	0.000184***	-85.03***
	(8.26)	(-25.06)
VOL	8.94E-14**	1.75E-7***
	(2.07)	(26.52)
VAR	1511***	25704913***
	(124.87)	(13.91)
PRC	-3.54E-6***	0.0387**
	(-27.49)	(1.97)
Medium Activity		
Intercept	0.000556***	15.94***
	(26.91)	(56.40)
IA	0.000359***	-18.38***
	(8.62)	(-32.33)
VOL	-3.31E-12***	2.72E-7***
	(-4.20)	(25.34)
VAR	1103***	1185498***
	(131.69)	(10.34)
PRC	-4.93E-6***	-0.0580***
	(-21.12)	(-18.15)
Low Activity		
Intercept	0.002542***	6.85***
	(37.87)	(50.88)
IA	0.000227***	-1.08***
	(2.58)	(-6.13)
VOL	-3.05E-10***	3.39E-7***
	(-23.63)	(13.09)
VAR	340.17***	82632***
	(84.13)	(10.19)
PRC	-0.000012***	-0.0596***
	(-11.38)	(-27.94)

*** and ** indicates significance at the 0.01 level and the 0.05 level respectively

Note: The numbers in parentheses are the t-statistics.

Table 19: Estimated Results for Equation (14.1) and Equation (14.2)

Instrumental Variable: Rank of IA		
Independent Variables	SPRD	DPTH
Intercept	0.00128***	18.71***
	(33.09)	(38.04)
IA	0.000363***	-16.22***
	(4.58)	(-16.09)
VOL	-4.34E-12***	2.57E-7***
	(-15.50)	(72.24)
VAR	404.49***	-66255.60**
	(158.52)	(-2.04)
PRC	-0.00001***	-0.1457***
	(-28.45)	(-25.14)
Adjusted R^2	0.7861	0.5197

*** and ** indicates significance at the 0.01 level and the 0.05 level respectively

Note: At the bottom the R^2 for Equation (14.1) and Equation (14.2) are both reported and the numbers in parentheses are the t-statistics.

Table 20: The Estimated Results for Equation (15.1) and Equation (15.2)

Independent Variables	SPRD	DPTH
Intercept	-0.00052***	17.59***
	(-24.41)	(42.33)
IA	0.000322***	-12.84***
	(8.97)	(-18.28)
VOL	1.55E-12***	2.62E-7***
	(8.94)	(76.96)
Volatility	2.27***	-424.26**
	(267.60)	(-2.56)
PRC	2.65E-6***	-0.1573***
	(8.86)	(-26.95)
Adjusted R^2	0.9098	0.5209

*** and ** indicates significance at the 0.01 level and the 0.05 level respectively

Note: At the bottom the R^2 for Equation (15.1) and Equation (15.2) are both reported and in parentheses are the t-statistics.

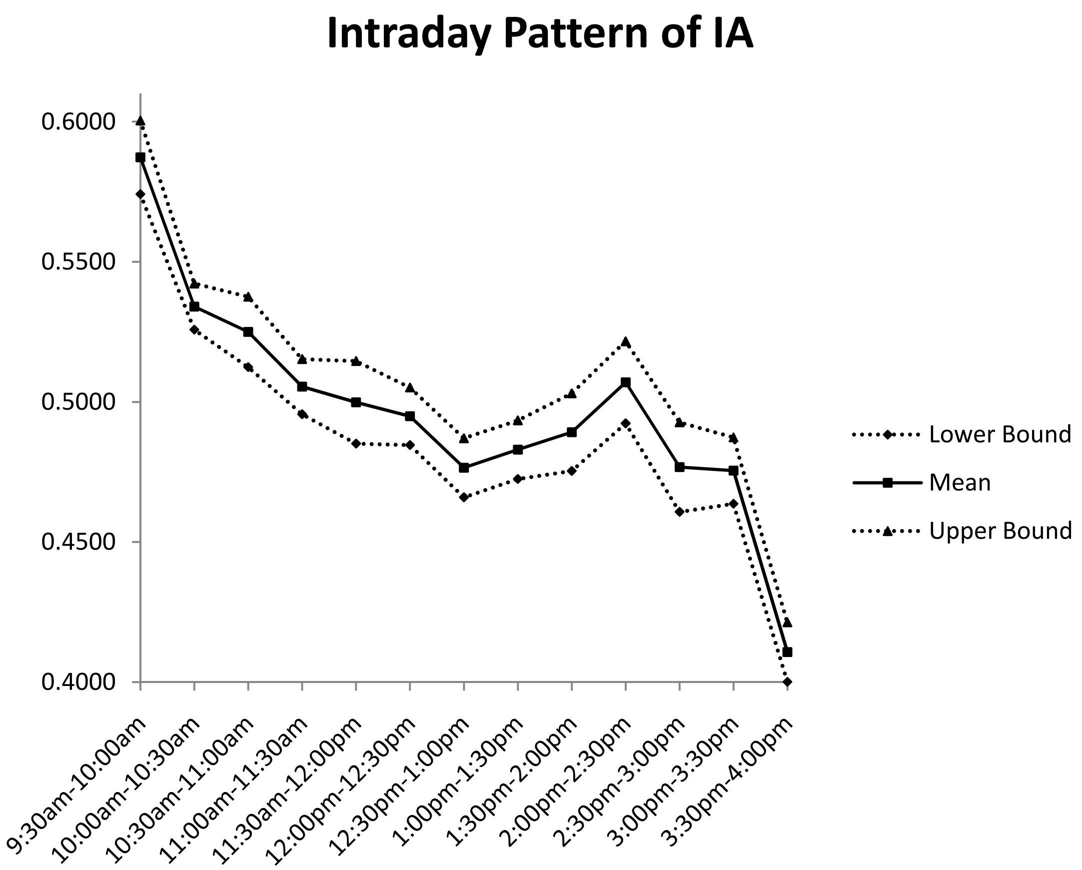

Figure 1a: Intraday Pattern of Information Asymmetry.
Note: The mean values and the lower and upper bound values of IA throughout the 13 trading interval obtained from Panel A in Table 4 are plotted here.

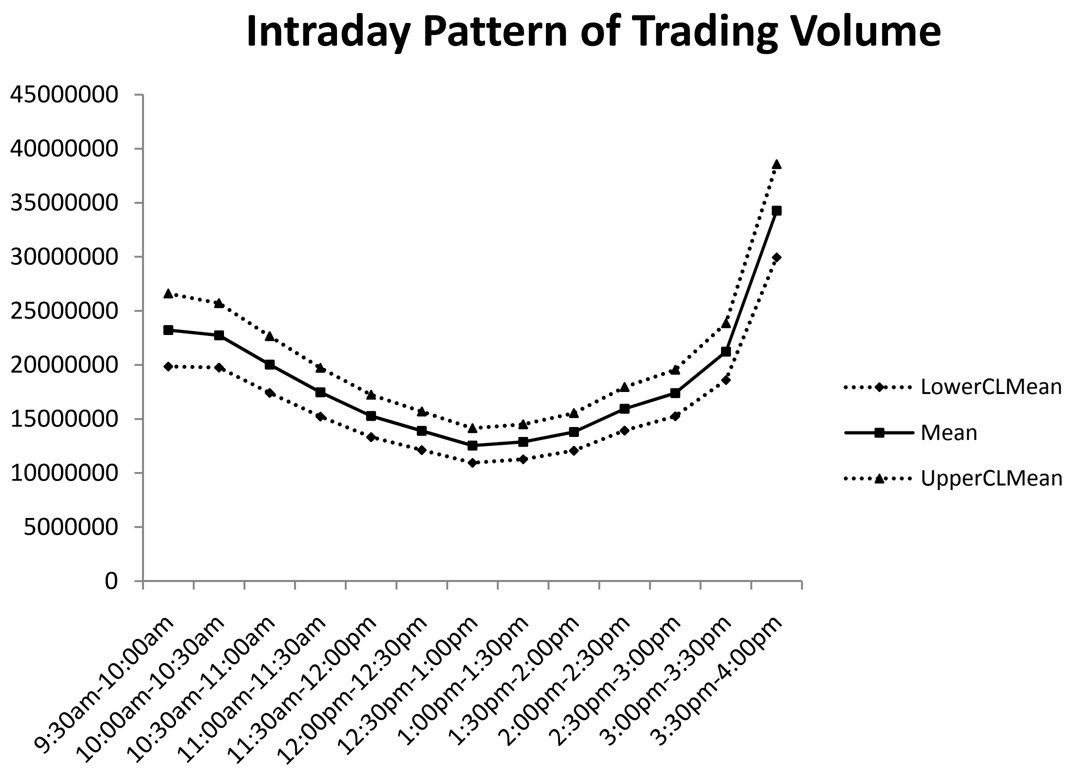

Figure 2b: Intraday Pattern of Trading Volume (Unit: Share).

Note 1: Here for each interval, the values for each marker are the mean trading volume for each of the 13 trading interval averaged across the 677 stocks. We also report the 95% confidence interval here.

Note 2: We also perform t-tests to compare weather the trading volume during noon period 12:30 pm to 1:30 pm is significantly lower than period before and after noon period.

T-tests:

12:30 pm to 1:30 versus 11:30 am to 12:30 pm: t-value = –17.10, p-value < 0.0001

12:30 pm to 1:30 versus 1:30 pm to 2:30 pm: t-value = –18.31, p-value < 0.0001

The t-tests indicate that the trading volume during noon period is significantly lower than period before and after noon period.

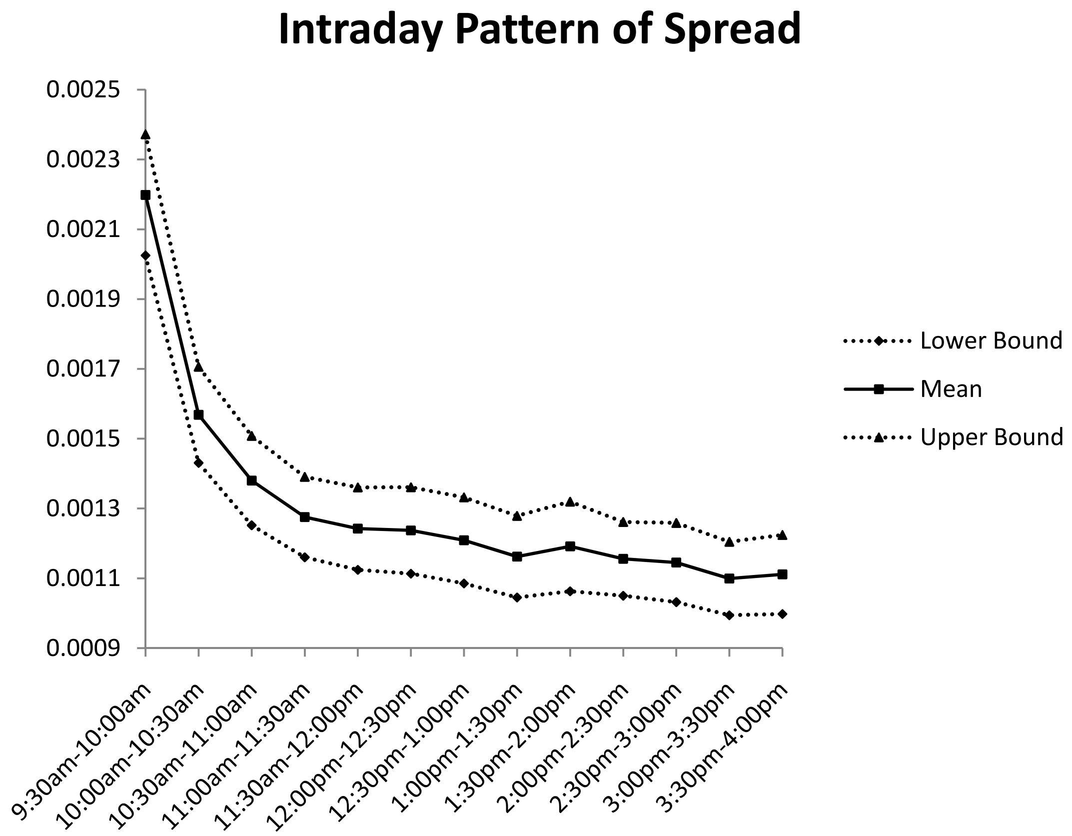

Figure 3: Intraday Pattern of Spread. (Unit: Dollar)
Note: The mean values and the lower and upper bound values of spread throughout the 13 trading interval obtained from Panel A in Table 5 are plotted here.

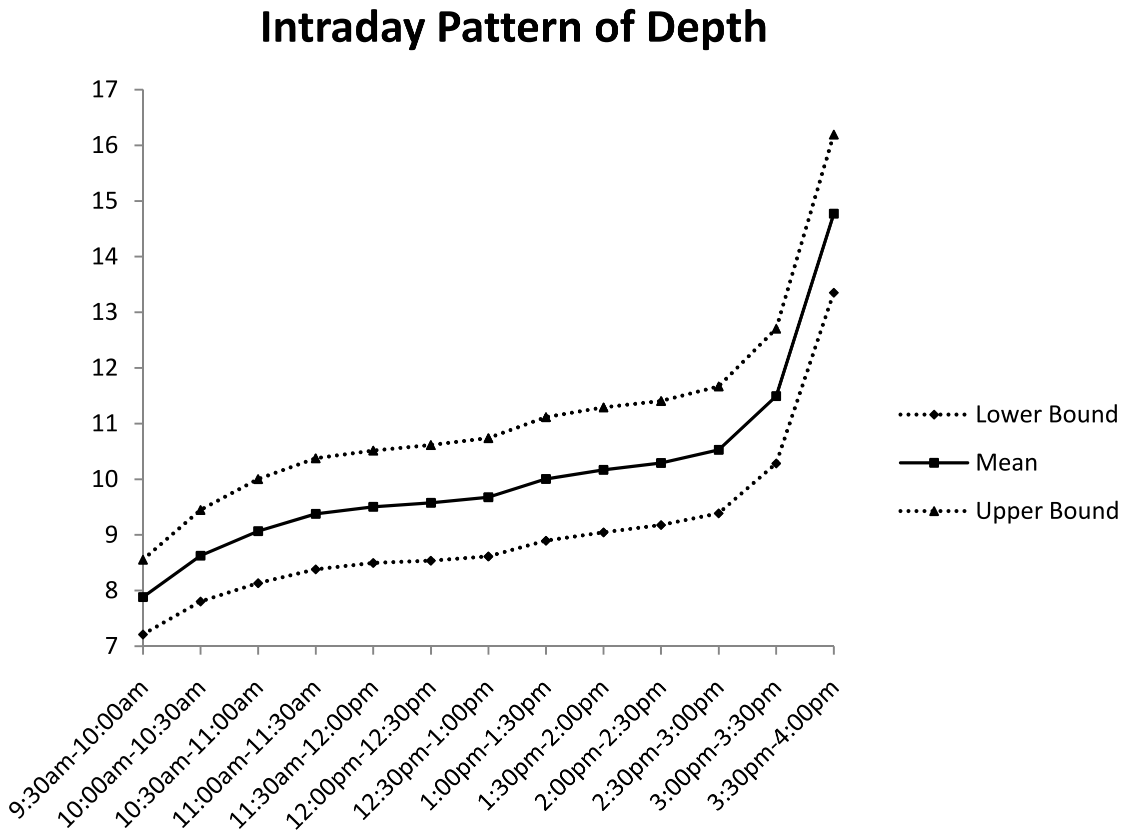

Figure 4: Intraday Pattern of Depth. (Unit: 100 shares)
The mean values and the lower and upper bound values of depth throughout the 13 trading interval obtained from Panel A in Table 6 are plotted here.

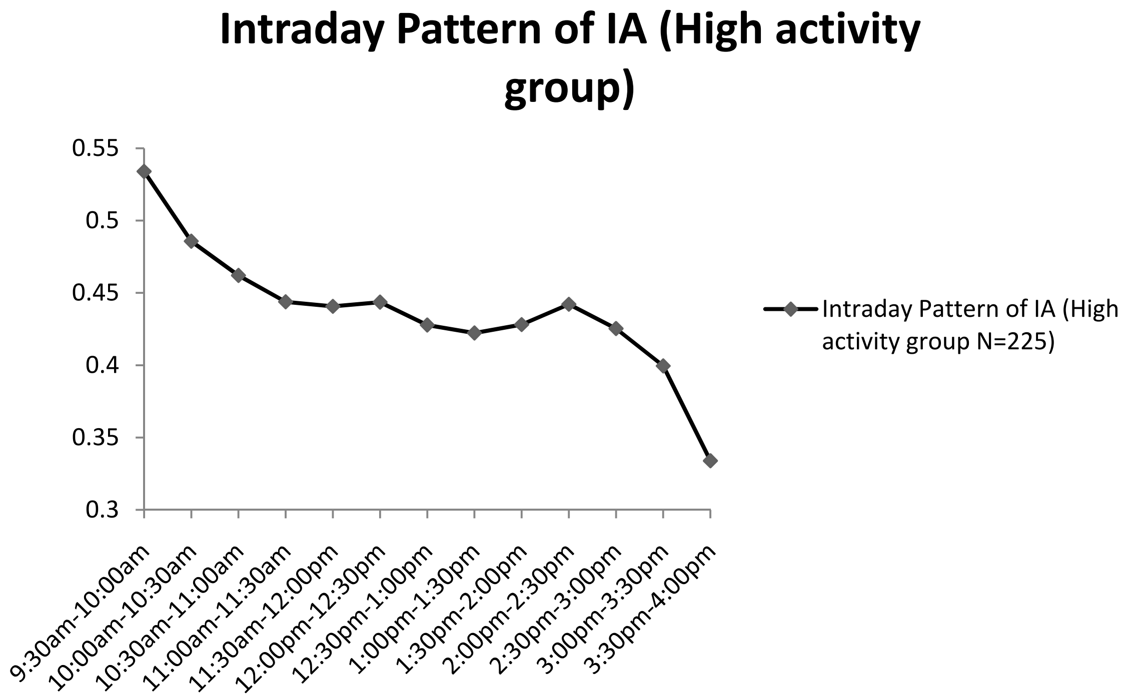

Figure 5a: The Intraday Pattern of Information Asymmetry for the High activity group of Stocks.

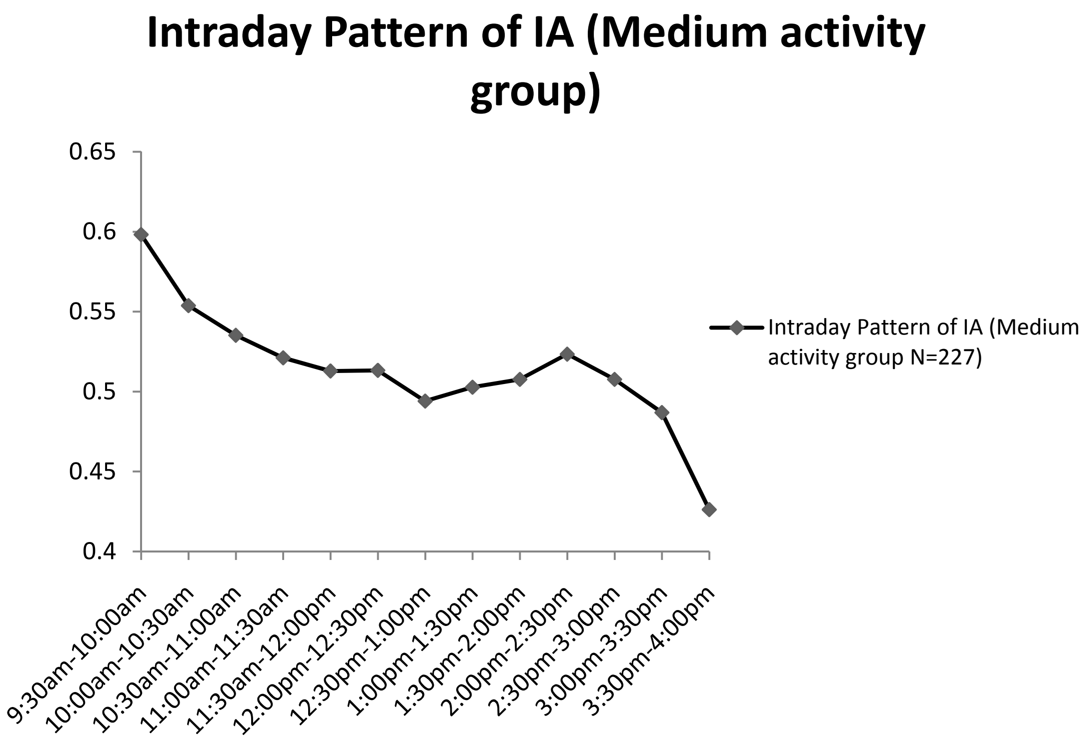

Figure 6b: The Intraday Pattern of Information Asymmetry for the Medium activity group of Stocks.

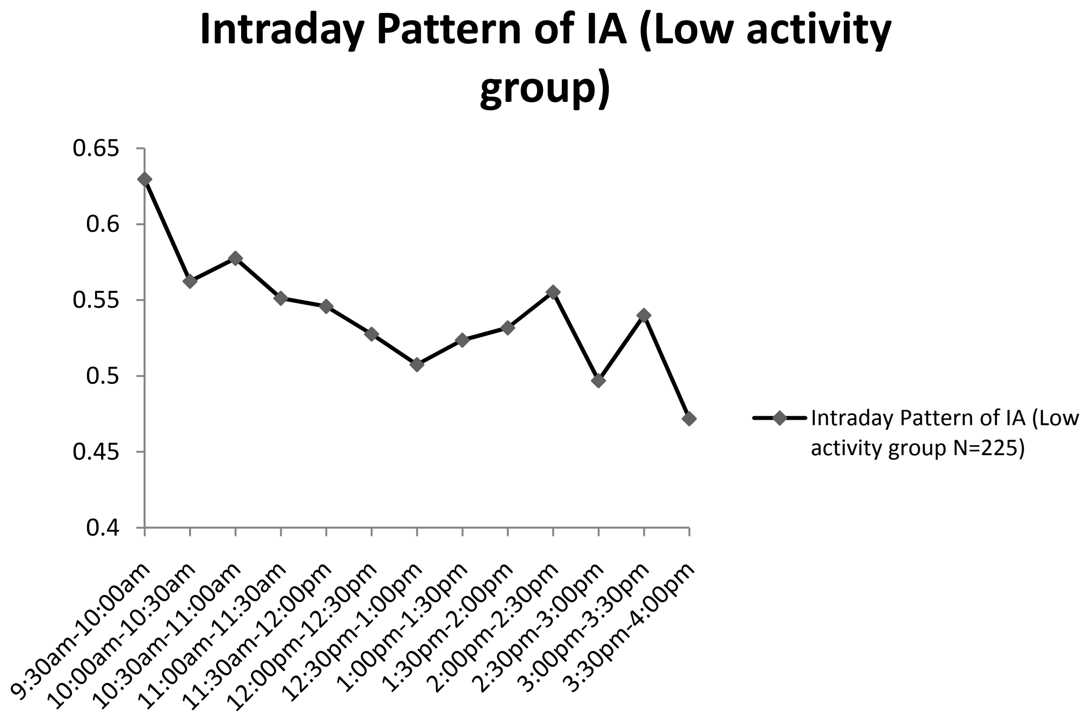

Figure 7c: The Intraday Pattern of Information Asymmetry for the Low activity group of Stocks.

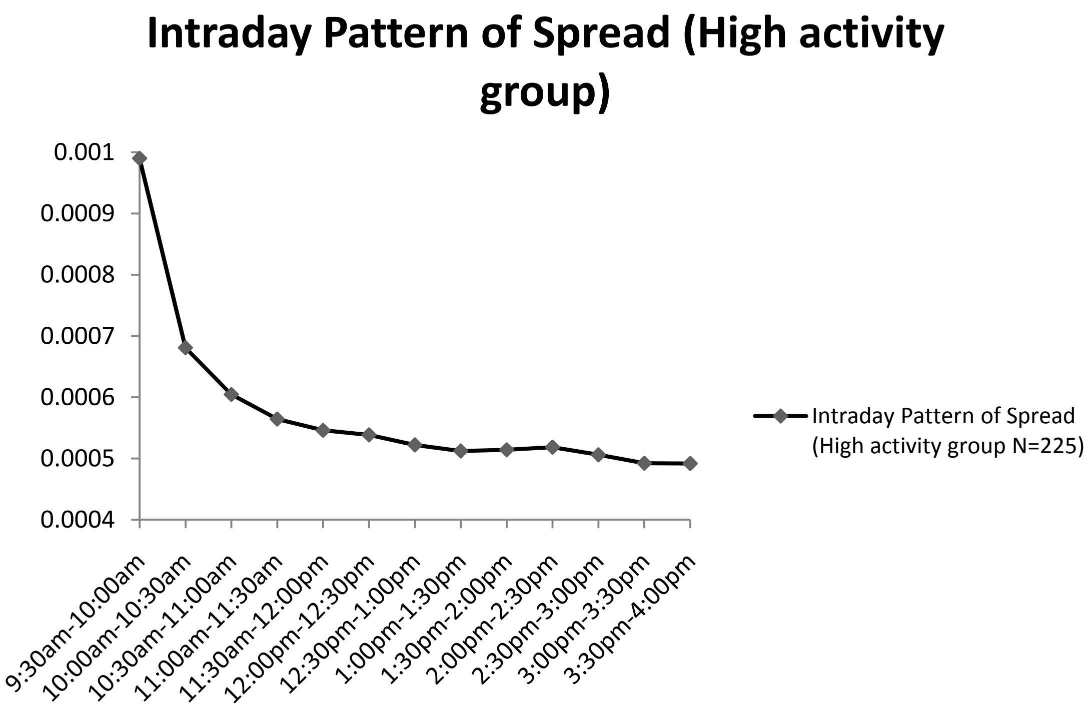

Figure 8a: The Intraday Pattern of Spread for the High activity group of Stocks. (Unit: dollar)

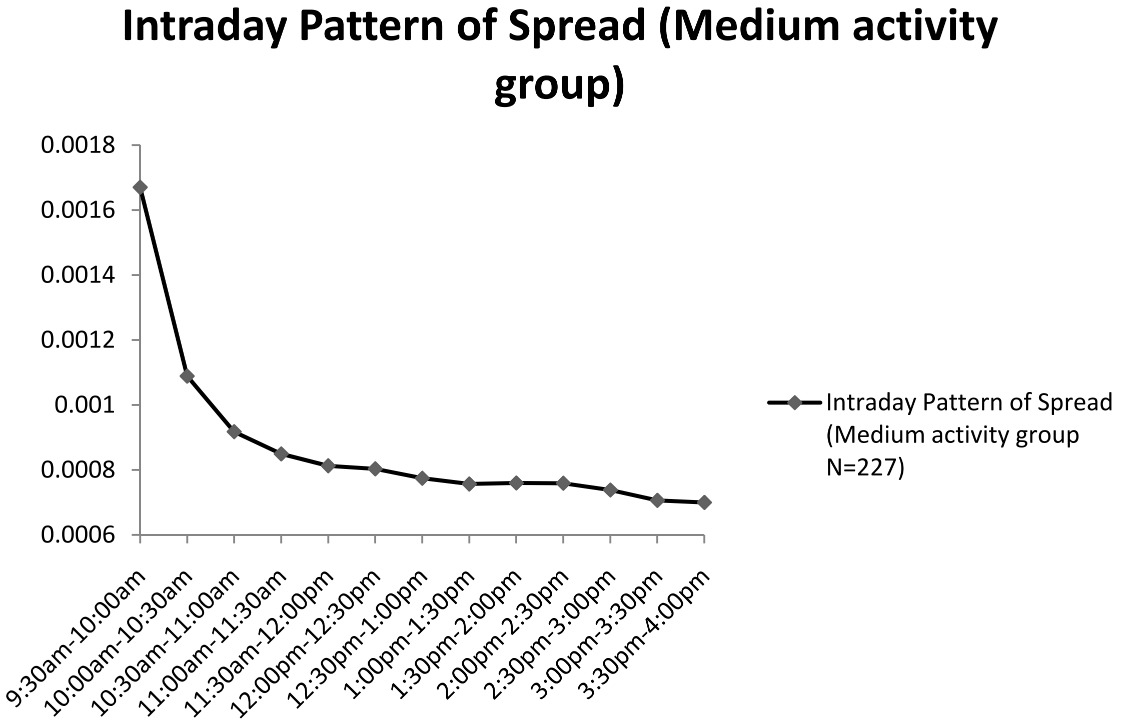

Figure 9b: The Intraday Pattern of Spread for the Medium activity group of Stocks. (Unit: dollar)

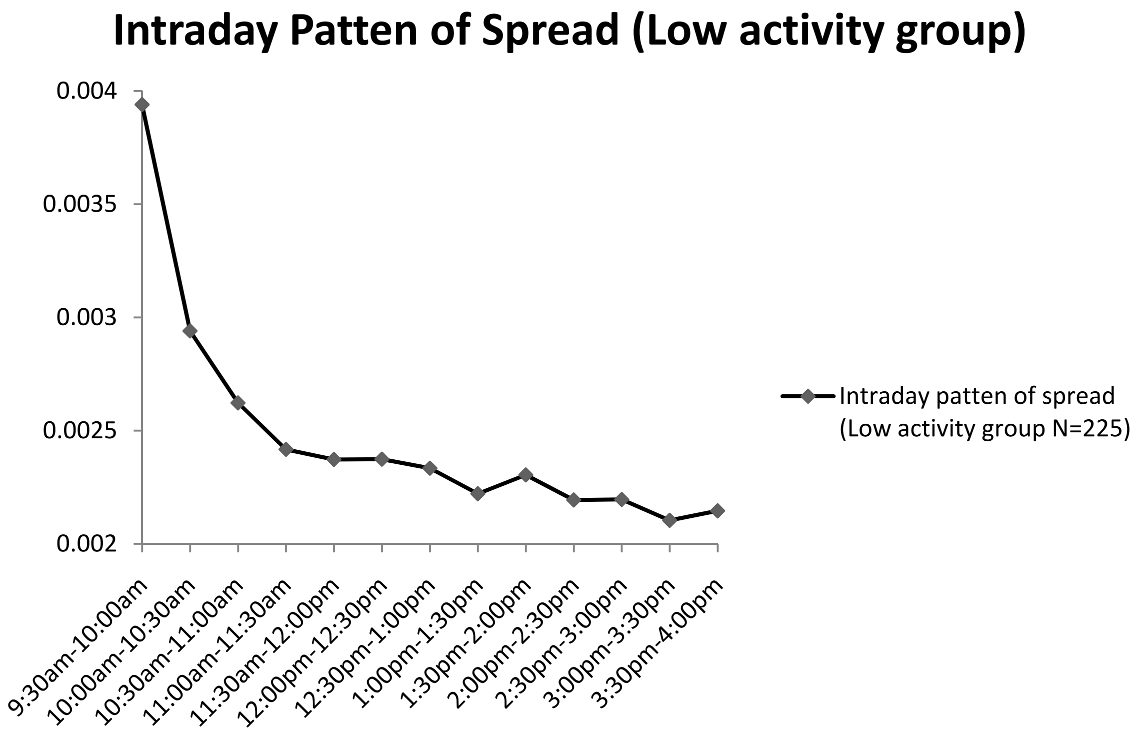

Figure 10c: The Intraday Pattern of Spread for the Low activity group of Stocks. (Unit: dollar)

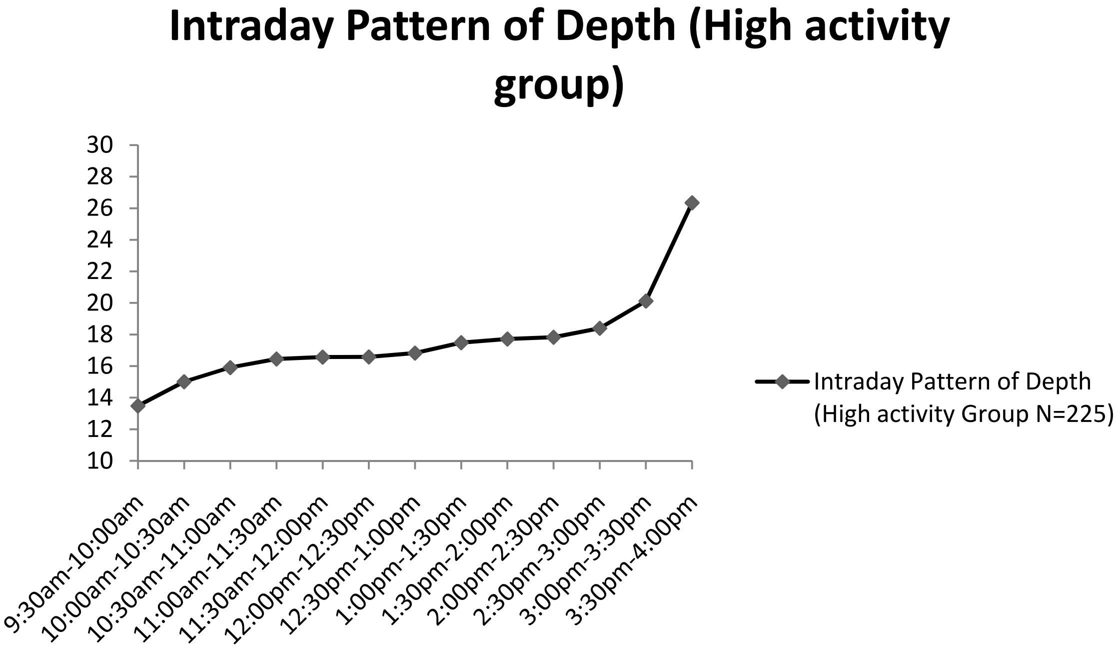

Figure 11a: The Intraday Pattern of Depth for the High activity group of Stocks. (Unit:100 shares)

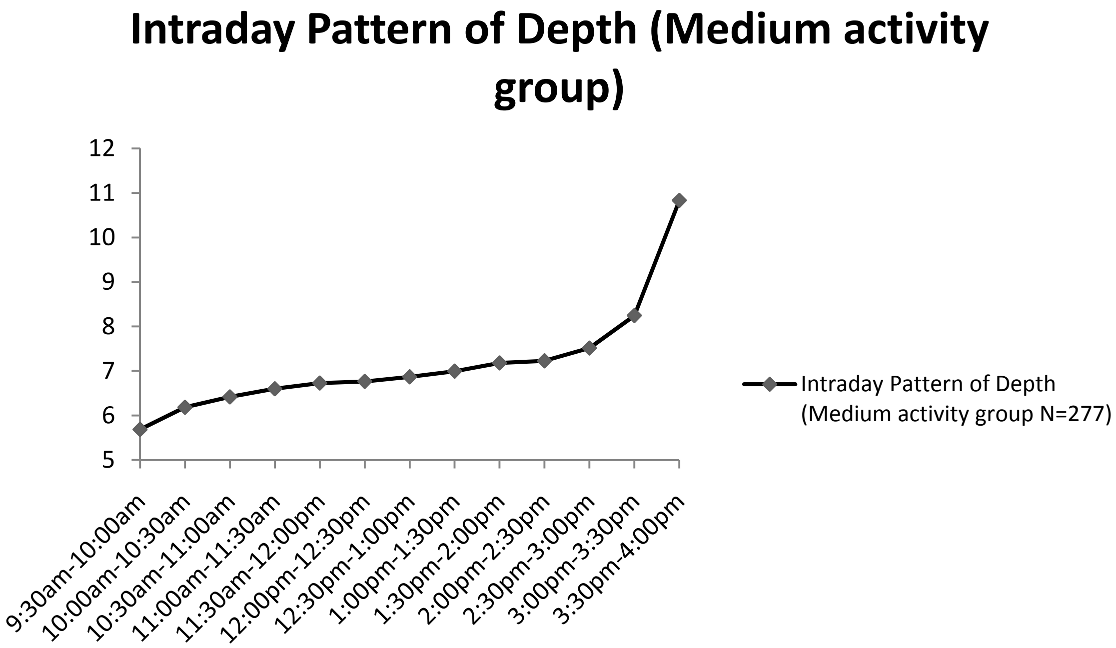

Figure 12b: The Intraday Pattern of Depth for the Medium activity group of Stocks. (Unit:100 shares)

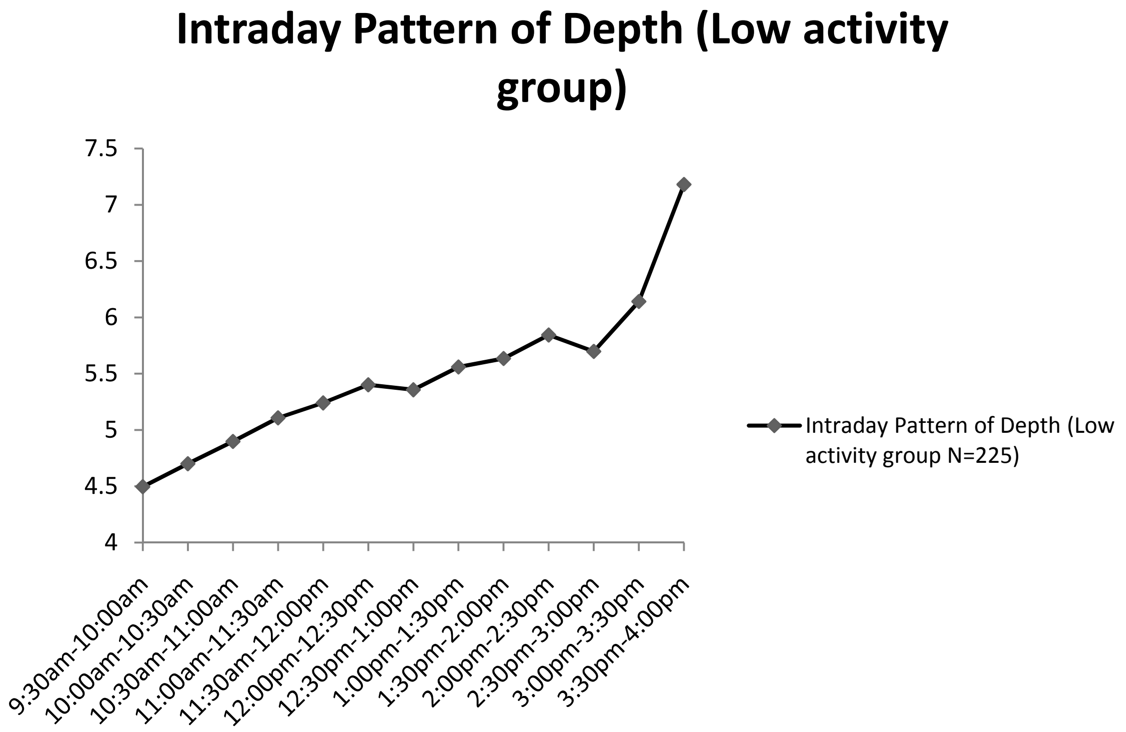

Figure 13c: The Intraday Pattern of Depth for the Low activity group of Stocks. (Units: 100 shares)

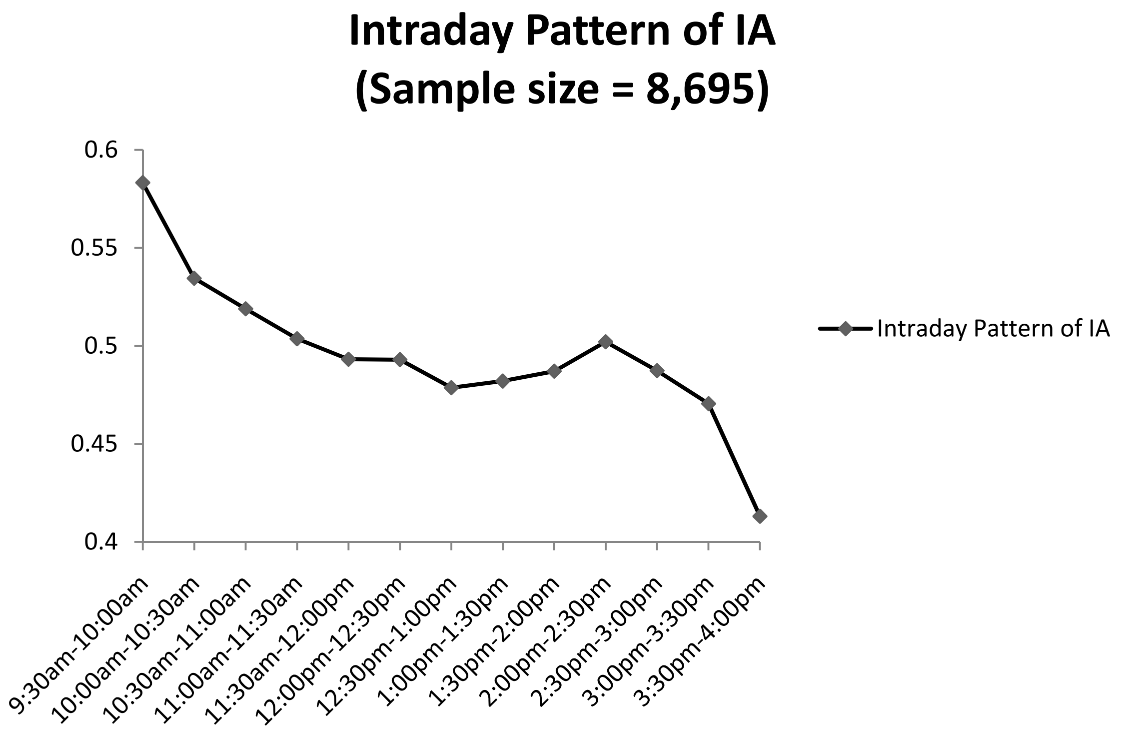

Figure 14: The Intraday Pattern of Information Asymmetry Based on the Regression Results of the 8,695 Equations.